Novel Ligands Targeting the DNA/RNA Hybrid and Telomeric Quadruplex as Potential Anticancer Agents

Islam, Mohammad Kaisarul

Awarding institution:
King's College London

Novel Ligands Targeting the DNA/RNA Hybrid and Telomeric Quadruplex as Potential Anticancer Agents

A dissertation submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Mohammad Kaisarul Islam
Institute of Pharmaceutical Sciences
School of Biomedical Sciences
King's College London

Supervisors

Professor David E. Thurston

Dr Khondaker Miraz Rahman

December 2015

Dedicated to my parents

Late Mohammad Fakhrul Islam

&

Mrs Momtaz Islam

Plagiarism statement

This thesis describes research conducted in the School of Pharmacy, UCL and King's College London, University of London between October 2011 and April 2015 under the supervision of Professor David E. Thurston and Dr Khondaker Miraz Rahman. I declare that the research described in this thesis is unique and I have clearly mentioned the part that was conducted by our research collaborators. I also declare that all the text in this thesis has been written by me and that the parts that have already appeared in publication have been suitably recorder.

Signature: _____ Date: _____

Acknowledgement

Firstly, I should like to express my sincere gratitude and profound thanks to my respected supervisors, Dr Khondaker M Rahman and Professor David E. Thurston for their valued supervision and encouraging support throughout my research work.

I appreciate the generous help and guidance of Dr Luengo Arratta Sandra of my group for her guidance, support and direction during the course of my chemistry laboratory work.

I am very thankful to Dr Paul J M Jackson for his continuous support in carrying out the molecular modelling studies of my synthesized molecules and subsequent advice regarding SAR of the molecules. I also acknowledge Mr Meir Touitou for his support in the molecular modelling and dynamics studies.

I should like to acknowledge Dr George Procopiou, Dr Nicolas Veillard and Dr Fabrizio Minicone for their valuable suggestions and guidance with the synthesis, and purification of the molecules.

I should like to express my gratitude to Dr Tam Bui from Biomolecular Spectroscopy Centre, Pharmaceutical Optical & Chiroptical Spectroscopy Facility, King's College London for helping with the CD analysis of the synthesised compounds.

I should also like to thank Amrit Varsha (visiting researcher) and Jeerapat Doungchawee (MSc student) for kind help in the biological evaluation of the synthesised molecules.

I am highly grateful to Kazi Sharmin Nahar, Kamrun Nahar, Christopher Chamberlain, Julia Mantaj, David Corcoran, AKM Azadur Rahman, Pietro Picconi, Rashedul Islam, Jennifer Auer, Paolo Andriollo and Fransesco Cascio from my group for their generous help and encouragement.

Finally, I want to acknowledge the tremendous support that I received from my family members, particularly my wife, Mrs Israt Zahan Eti, my mother, Mrs Momtaz Islam, my uncle, Dr M Khairul Hossain and my aunty, Mrs Sabera Hossain during the course of my PhD studies.

Table of Contents

List of Figures **10**

List of Tables **18**

Abbreviations **19**

Abstract **21**

Chapter 1: Introduction **23**

1.1 Cancer and General Overview **23**

1.2 Hallmarks of Cancer **26**

1.2.1 Sustaining proliferative signalling **26**

1.2.2 Evading growth suppressors **27**

1.2.3 Evading programmed cell death or apoptosis **27**

1.2.4 Limitless replicative potential **28**

1.2.5 Sustained angiogenesis **28**

1.2.6 Tissue invasion and metastasis **28**

1.3 Treatment Options for Cancer **30**

1.3.1 Chemotherapy **30**

1.3.1.1 DNA targeted agents 31

1.3.1.2 Antimetabolites 34

1.3.1.3 Antitubulin Agents 35

1.3.1.4 Molecularly targeted agents 35

1.3.1.5 Hormonal therapy 37

1.3.2 Antibody-drug conjugates (ADC) for cancer therapy **38**

1.3.3 Radiotherapy **38**

1.3.4 Biotherapy or immunotherapy **39**

1.3.5 Photodynamic therapy (PDT) **39**

1.3.6 Surgery **39**

1.4 DNA Structure **39**

1.4.1 Double Helical DNA **42**

1.4.1.1 A-DNA 43

1.4.1.2 B-DNA 44

1.4.1.3 Z-DNA 44

1.4.2 Triplex DNA **46**
1.4.3 Quadruplex DNA **47**
1.4.4 i-Motif **49**
1.5 DNA Major and Minor Grooves **50**
1.6 Telomeric DNA **51**
1.6.1 Functional Assembly of Human Telomerase **55**
1.6.1.1 RNA Subunit, hTR 55
1.6.1.2 Protein Subunit, hTERT 56
1.6.2 Telomeric DNA/RNA Hybrid Duplex **56**
1.6.2.1 Structure and stability of DNA/RNA hybrid 59
1.6.3 Telomeric G-quadruplex Structure **61**
1.6.3.1 G-quadruplex structures in promoter regions 68
1.7 Telomerase and Cancer **69**
1.7.1 Mechanisms of Telomerase Up-Regulation in Cancer **70**
1.7.2 Telomere Lengthening by ALT **72**
1.8 Telomerase and the Aging Process **72**
1.9 Opportunities for Telomerase Targeted Therapies **74**
1.9.1 Telomeric DNA/RNA Hybrid Duplex as a Target in Cancer **76**
1.9.2 Telomeric G-quadruplex as a target **80**
1.9.3 Antisense oligonucleotides **89**
1.9.4 Targeting telomere and telomerase-associated proteins **89**
1.9.5 Small molecule inhibitors of hTERT **91**
1.9.6 Telomerase based immunotherapies **91**
1.10 Telomerase inhibition measurement **92**
1.10.1 Detection of telomerase activity and expression **93**
1.10.1.1 Telomeric Repeat Amplification Protocol (TRAP) Assay 93
1.10.1.2 Direct detection of telomerase-synthesized DNA 95
1.11 Design and Rationale of this Project **96**
1.11.1 Background **96**
1.11.2 Rationale of the Project **97**
1.11.3 Aims and Objectives **99**

Chapter 2: Methods and Materials **101**

2.1 Sourcing of Chemicals **101**

2.2 Analytical Tools **101**

2.3 Purification Techniques **102**

2.3.1 Flash Column Chromatographic Technique **102**

2.3.2 Catch and Release Method **102**

2.3.3 Trituration Technique **103**

2.4 General Amide Coupling Reaction **103**

2.5 Methods Used for Biophysical and Biological Evaluation **105**

2.5.1 Fluorescence Resonance Energy Transfer (FRET) Assay **105**

2.5.1.1 Procedure involved in FRET Assay 111

2.5.1.2 Melting curve analysis 113

2.5.1.3 Limitation of FRET melting assay 114

2.5.2 Molecular Modelling and Molecular Dynamics (MD) studies **115**

2.5.2.1 Process involved in molecular modelling experimentation for DNA/RNA hybrid duplex 116

2.5.2.2 Process involved in molecular modelling experimentation for hTelo Quadruplex sequence 117

2.5.2.3 Limitations of MD simulation studies 118

2.5.3 Circular Dichroism (CD) Spectroscopy **119**

2.5.3.1 Procedure involved in CD Analysis 123

2.5.3.2 Limitations of CD spectroscopic analysis 125

2.5.4 Cytotoxicity Test with MTT Assay **125**

2.5.4.1 Procedure involved in MTT Assay 127

2.5.4.2 Limitations of MTT assay 130

2.5.5 Preliminary Biological Evaluation **131**

2.5.5.1 Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) Assay 131

2.5.5.2 Electrophoretic Mobility Shift Assay (EMSA) 134

Chapter 3A: Chemistry: Results and Discussion **136**

3A.1 Introduction **136**

3A.2 Molecules in Library-1 ..... **139**

3A.2.1 Design Rationale of Library-1 Molecules **139**

3A.2.2 Synthesis of bis-benzimidazole molecules in library-1 140
3A.3 Molecules in Library-2 144
3A.3.1 Design rationale of library-2 molecules 144
3A.3.2 Synthesis of bis-quinoline type molecules in library-2A 145
3A.3.3 Synthesis of bis-(6-fluoro-2-methylquinoline) molecules in library-2B 149
3A.3.4 Synthesis of bis-indole molecules in library-2C 152
3A.4 Molecules in Library-3 155
3A.4.1 Design Rationale of Library-3 Molecules 155
3A.4.2 Synthesis of bis-acridine type molecules in library-3A 156
3A.4.3 Synthesis of bis- (2-fluoro-7,8,9,10-tetrahydro-6H-cyclohepta[b] quinoline-11-carboxamide in library-3B 158
3A.5 Molecules in Library-4 161
3A.5.1 Design Rationale of Library-4 Molecules 161
3A.5.2 Synthesis of bis-benzofuran molecules in library-4 162
Chapter 3B: Biophysical and Biological Evaluation: Results and Discussion 169
3B.1 Fluorescence Resonance Energy Transfer (FRET) Assay 169
3B.2 Molecular Modelling and Molecular Dynamics (MD) Studies 175
3B.3 Circular Dichroism (CD) Analysis 191
3B.4 Cytotoxicity Test Using the MTT Assay 204
3B.4.1 MTT Assay Results with MDA-MB-231 Cell Line (Breast Cancer Cell Line) 205
3B.4.2 MTT Assay Results with HeLa Cell Line (Cervical Cancer Cell Line) 207
3B.4.3 MTT Assay Results with NCI H1975 Cell Line (Non-Small Lung Cancer Cell Line) 209
3B.5 Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) Assay 210
3B.5.1 hTERT Expression 212
3B.5.2 Cyclin B Expression 213
3B.6 Electrophoretic Mobility Shift Assay 215
Chapter 4: Conclusion and Future Work 217
Appendix 1: Experimental, Synthesis and Characterisation of Molecules 220
A1.1 Synthesis of bis-benzo[d]imidazole type molecules in library-1 220
A1.2 Synthesis of bis-quinoline type molecules in library-2A 225
A1.3 Synthesis of bis-(6-fluoro-2-methylquinoline) molecules in library-2B 229

A1.4 Synthesis of bis-indole compounds in library-2C 233

A1.5 Synthesis of bis-acridine molecules in library-3A..... 236

A1.6 Synthesis of bis- (2-fluoro-7,8,9,10-tetrahydro-6H-cyclohepta[b] quinoline-11-carboxamide in library-3B..... 238

A1.7 Synthesis of bis-benzofuran type molecules in library-4 242

Appendix 2: Experimental data and Graphs 259

A2.1 FRET Melting Result..... 259

A2.1.1 DNA/RNA Hybrid duplex..... 259

A2.1.2 Telomeric quadruplex (F21T) structure 262

A2.2 UV calibration curves obtained on Circular Dichroism (CD) Studies 264

A2.2.1 DNA/RNA Hybrid duplex..... 265

A2.2.2 Telomeric quadruplex (F21T) structure 266

A2.3 IC_{50} curves observed in Cytotoxicity Studies (MTT Assay) 267

A2.3.1 MDA-MB-231 Cell Line (Breast Cancer Cell Line)..... 267

A2.3.2 HeLa Cell Line (Cervical Cancer Cell Line)..... 269

A2.3.2 NCI H1975 Cell Line (Non-Small Lung Cancer Cell Line)..... 271

References 274

List of Figures

Chapter 1		
Figure 1.1	Cancer cells dispersion mechanism (*i.e.*, metastasis)	24
Figure 1.2	Reproduced illustration of six hallmarks of cancer and two emerging hallmarks of cancer, adapted from Hanahan and co-workers	26
Figure 1.3	(a) A timeline for the discovery and evolution of DNA-interactive agents. (b) Illustrates the types of modifications that can form on DNA (from left to right, crosslinks, intercalation, DNA-strand cleavage and code-reading molecules). The DNA helix is also associated with proteins (e.g. topoisomerase II) and secondary DNA structures (*e.g*. G-quadruplex). (c) Detailed illustration of the different types of DNA modification. (Adapted from DNA and its associated processes as targets for cancer therapy by Laurence H. Hurley)	31
Figure 1.4	Structures of some DNA interacting agents	33
Figure 1.5	The structure of quarfloxin (1.12) and its mechanism of action	34
Figure 1.6	Structural modification between dihydrofolic acid and methotrexate (circled in red)	34
Figure 1.7	Structure of vincristine and paclitaxel	35
Figure 1.8	Components of DNA; the four common deoxyribonucleotides connected by phosphodiester bonds to form a single-strand (5′→3′)	41
Figure 1.9	(a) Complementary base pairs in the DNA double helix, (b) double stranded DNA helical structure	43
Figure 1.10	Three major types of DNA double helix, (a) A-DNA, (b) B-DNA and (c) Z-DNA	44
Figure 1.11	The helix axis of A-, B-, and Z-DNA	45
Figure 1.12	Diagram of triplex DNA structure	46
Figure 1.13	(a) G-quartet with central stabilizing cation, (b) Potential topologies for G-quadruplex structures, topologies constructed from four parallel strands (i), from two strands that are non-crossing (ii), cross-over (iii) and from a single strand (iv)	48
Figure 1.14	(a) Hemi-protonated C-C^+ pairs (note*, hydrogen bonds are not of an appropriate length*) (b) Intermolecular i-motif pattern.	49
Figure 1.15	(a) Major and minor groove formation at a dC-dG base pair and (b) major and minor groove formation at a dT-dA base pair	50
Figure 1.16	The consequences that arise from telomere function and dysfunction	51
Figure 1.17	Schematic representation of the end replication problem	52
Figure 1.18	Telomere structure: (A) Telomeric DNA consists of repetitive DNA sequence of a duplex region and single-stranded DNA G-strand	53

overhang. (B) The shelterin complex binds to both the duplex and single-stranded DNA regions through specific protein/DNA interactions. (C) Formation of t-loop involving strand invasion of the G-overhang to create a displacement-loop (D-loop)

Figure 1.19 Schematic representation of a functioning model for telomerase activity 58

Figure 1.20 Probable places of the G-quadruplex structures within cells: in nucleus (a) during replication (b) transient single stranded form of DNA (c) telomeric single stranded G-rich overhangs; outside nucleus (d) G-quadruplexes in mRNA. (G-quadruplex mediated blockade during replication, transcription and translation labelled by red cross) 62

Figure 1.21 Telomere length regulation (a) Inhibition of the telomerase enzyme by G-quadruplex formation at the 3′ end which was further stabilized by ligand interaction inhibiting further telomere elongation. (b) Diagram showing the effects of inhibition of telomere elongation; critically formed hTelo that ultimately leads to senescence/apoptosis 63

Figure 1.22 Illustration of three possible G-quadruplex structures that may exist together as a mixture in solution, where (a) tetramer RNA-quadruplex can form by 6-mer hTelo RNA sequence, (b) dimeric DNA-quadruplex can form by 16-mer hTelo DNA sequence and (c) DNA/RNA hybrid-quadruplex can form by 6-mer RNA and 16-mer hTelo DNA 64

Figure 1.23 Different forms of intramolecular quadruplex formation by the hTelo repeat sequence (d[AGGG(TTAGGG)$_3$]), where loops are coloured red, *anti*- and *syn*-guanines are coloured cyan and magenta, respectively and W, M and N denote wide, medium and narrow grooves, respectively 65

Figure 1.24 Folding patterns of the hTelo sequence (22-mer); (A) propeller-type parallel-stranded intramolecular G-quadruplex formed in K^+ solution, (B) basket-type mixed parallel/antiparallel-stranded intramolecular G-quadruplex in Na^+ solution 66

Figure 1.25 The crystal structure of the bimolecular quadruplex formed by the *Oxythricha nova* telomeric sequence d($G_4T_4G_4$) (PDB entry 1JPQ); (a) overall topology is indicated by the orange ribbon with purple spheres of K^+ ions, and (b) a projection down the central passage representing the relative widths of the four grooves 66

Figure 1.26 Structures of the hTelo quadruplex for the d[AGGG(TTAGGG)$_3$] sequence, (a) in Na^+ solution with one diagonal and two lateral loops (PDB entry 143D), (b) in K^+ solution that forms A-DNA with three strand-reversal loops (PDB entry 1KF1), (c) in K^+ solution, that forms B-DNA with one strand-reversal and two lateral loops, as determined by NMR 67

Figure 1.27 (a) Distinct structures of hybrid-1 and hybrid-2, established by NMR studies of telomeric G-quadruplexes (here, guanine, adenine and 68

thymine are in the yellow, red and blue blocks). (b) DNA secondary structure model formed due to compact-stacking of hybrid-type G-quadruplexes in an equilibrium of hybrid-1 and hybrid-2 in K^+ solution

Figure 1.28 Summary of telomerase activity in humans, where tumour samples were evaluated in a TRAP assay (telomerase positive samples are compared with matched control tissue and denoted in %) 70

Figure 1.29 Telomerase up-regulation in cancer cells can occur by a variety of mechanisms 71

Figure 1.30 Multiple deregulation steps associated progression of the ALT mechanism 72

Figure 1.31 Scheme of factors affecting the telomere length in human tissues that also contribute to the aging process 74

Figure 1.32 The action of telomerase may be inhibited at multiple points in its "life cycle" 75

Figure 1.33 Possible ligand binding mechanisms to inhibit telomerase activity by stabilizing the DNA/RNA hybrid duplex, as proposed by West and co-workers 77

Figure 1.34 Some reported molecules that showed binding affinity towards a DNA/RNA hybrid duplex (poly(rA):poly(dT)) sequence 78

Figure 1.35 Structure of Paromomycin 79

Figure 1.36 Structure of neomycin (1.31) and neomycin-methidium conjugate (1.32) 79

Figure 1.37 Structure of a novel diaryl substituent and pyrimidine analogue with a selective affinity for DNA/RNA hybrid duplexes 80

Figure 1.38 (a) Metal ion is shown coordinated between eight carbonyl O6 with an average distance of 2.73 Å; (b) structural illustration of O6-NH bonding for each quartet (*note, hydrogen bonds are not of an appropriate length*) 81

Figure 1.39 Roles of telomeric DNA and telomerase; G-quadruplex formation can inhibit telomerase activity 82

Figure 1.40 2,6-Diamidoanthraquinone analogue 83

Figure 1.41 Example of fluorenone and acridine molecules 84

Figure 1.42 Example of triazine type molecules 85

Figure 1.43 Example of cyclic porphyrines and five-ring acridine 86

Figure 1.44 Example of naphthyridine dimer and cyclic naphthalene structure 87

Figure 1.45 Examples of reported non-cyclic aromatic structure with planar shape and a positively charged moiety 88

Figure 1.46 FDA approved poly-ADP ribose polymerase (PARP) inhibitor 90

Figure 1.47 Structure of BIBR1532 and its analogue, small molecule targeting the hTERT 91

Figure 1.48	Summary of the original TRAP assay	94
Figure 1.49	Structural features of NCI library compound NSC 273829	97
Figure 1.50	Expected outcome from this project targeting DNA/RNA hybrid duplexes to inhibit telomerase activity within cancer cells	99

Chapter 2

Figure 2.1	Procedure for the 'Catch and Release' purification method	103
Figure 2.2	Schematic diagram of the HOBt/DIC-mediated amide coupling reaction	104
Figure 2.3	Principle of the FRET mechanism, the 'donor' fluorophore (Do) absorbs a photon and transfers this energy by a non-radiative process to the 'acceptor' fluorophore (Ac)	105
Figure 2.4	Jablonski diagram illustrating the coupled transitions involved between donor emission and acceptor absorbance in FRET	108
Figure 2.5	FRET transfer mechanism and application to a ligand on a DNA duplex sequence	109
Figure 2.6	Structure of the dual probe-labelled quadruplex, stabilized by a ligand (oval shaped) and subsequent schematic representation of FRET-based experiments	109
Figure 2.7	Structure of FAM (2.1) and TAMRA (2.2)	110
Figure 2.8	(a) Hairpin structure of telomeric DNA/RNA hybrid duplex (red indicates the RNA nucleotides) and (b) Hairpin structure of control cDRH duplex	111
Figure 2.9	Typical dissociation curve of DNA melting	114
Figure 2.10	CD spectrum of RNA/DNA hybrid (*i.e.*, poly(rA)-poly(dT)) in BPES buffer	119
Figure 2.11	Conversion of CD spectra due to formation of hTelo DNA/complementary RNA sequence-assisted hybrid quadruplexes, in presence of both monovalent cation Na^+ and K^+, as reported by Xu and colleagues	120
Figure 2.12	(a) Molecular arrangement and outlined map of parallel quadruplex structure (b) CD spectra reported by Xiaoyan and co-workers for A_o = TTAGGG, B_o = $(TTAGGG)_2$ and C_o = $(TTAGGG)_4$ in Tris-HCl buffer without monovalent metal ions at 20°C, pH 7.5	121
Figure 2.13	CD spectra of TTAGGG (A), $(TTAGGG)_2$ (B) and $(TTAGGG)_4$ (C) in Tris-HCl buffer at 20 °C, pH 7.5 in presence of (a) 80 mrnol/L Na^+ in solution (b) 80 mmol/L K^+ in solution, as reported by Xiaoyan and co-workers	123
Figure 2.14	Schematic diagram illustrating the conversion of MTT to formazan (purple) in the presence of an electron acceptor, 1-methoxy-phenazine	127

Figure 2.15	MTT assay plate layout in general	129
Figure 2.16	Schematic representation of the steps involved in the RT-PCR process	132
Figure 2.17	General principle of the electrophoretic mobility shift assay	134
Chapter 3A		
Figure 3A.1	Experimental approach to develop small molecule targeting telomeric DNA/RNA hybrid duplex	137
Figure 3A.2	General structure of the library-1 molecules	138
Figure 3A.3	General structures of other library molecules synthesized in this project	138
Figure 3A.4	The benzimidazole building block used in the synthesis of library-1 molecules	139
Figure 3A.5	Comparison between library-1 molecules with NSC 273829	140
Figure 3A.6	General synthetic scheme for library-1 molecules	141
Figure 3A.7	Structural variation from library-1 molecules to library-2A, -2B and -2C molecules and features in common with NSC 273829	145
Figure 3A.8	General synthetic scheme for library-2A molecules	146
Figure 3A.9	General synthetic scheme for library-2B molecules	149
Figure 3A.10	General synthetic scheme for library-2C molecules	153
Figure 3A.11	Effect of the insertion of cyclohexyl and cycloheptyl ring in library-3A and -3B; 3D view shows the distortion in molecule's planarity	155
Figure 3A.12	General synthetic scheme for library-3A molecules	156
Figure 3A.13	General synthetic scheme for library-3B molecules	158
Figure 3A.14	Structural features of library-4 molecules	162
Figure 3A.15	Main building blocks used in the synthesis of library-4 molecules	163
Figure 3A.16	Monoamide coupled product (1st step)	163
Figure 3A.17	Hydrogenated monoamide coupled product (2nd step)	164
Figure 3A.18	General scheme of reaction involved in the synthesis of library-4 molecules	165
Chapter 3B		
Figure 3B.1	Structural features and highlighted modification in library-1 molecules	170
Figure 3B.2	FRET melting curve for compound 3.6 at three concentrations (5 μM, 2 μM and 1 μM), where (a) change of intensity of fluorescence and over the change of temperature (-dI/dT) and increment of fluorescence due to increase of temperature	171
Figure 3B.3	Molecular model of the DNA/RNA hybrid duplex (5′-TTA-GGG-TTA-GGG-TTT-TTT-CCC-UAA-CCC-UAA-3′). The RNA strand is	176

represented in red and the DNA strand in grey. DNA and RNA bases are indicated in stick form

Figure 3B.4 Snapshot of a 10ns implicit solvent molecular dynamics simulation of compound 3.6 (blue spheres) interacting with DRH. Both benzimidazole moieties of the molecule intercalate between base pairs; one between G11:C20 and G10:C21 (green) and the second between A23:T8 and A24:T7 (both yellow) 177

Figure 3B.5 Snapshot of a 10ns explicit solvent molecular dynamics simulation of compound 3.6 (blue and yellow spheres) interacting with the control duplex sequence, cDRH. Both benzimidazole portions (yellow) point away from the DNA groove and do not interact with the cDRH 178

Figure 3B.6 Snapshot of a 10ns implicit solvent molecular dynamics simulation of 3.8 (blue spheres) with DRH. Both benzimidazole moieties of the molecule intercalate the sequence; one between G4:C27 and G5:C26 (green) and the second between the A9 and A10 nucleobases (yellow) 179

Figure 3B.7 Snapshot of a 10ns explicit solvent molecular dynamics simulation of NSC 273829 (green spheres) interacting with the DRH sequence (top, grey and red) and the cDRH (bottom, grey). The molecule is accommodated in each, but the shape-fit of the molecule to the DRH is more favourable due to the different topological features of the DRH structure 180

Figure 3B.8 Snapshot of a 10ns molecular dynamics simulation of compound 3.39 (green) in the minor groove of the DRH structure. The molecule does not intercalate into the DRH structure and fails to form any significant bonding 182

Figure 3B.9 Snapshot of a 10ns molecular dynamics simulation of compound 3.22 (yellow) seated in the minor groove of the DRH structure. The molecule does not intercalate within the DNA and fails to form any significant bonding interactions with the DRH structure 183

Figure 3B.10 Snapshot of a 10ns molecular dynamics simulation of compound 3.10 (blue) seated in the minor groove of the DRH structure. The molecule part-intercalates into the DRH structure but fails to form significant bonding interactions with the DRH structure 184

Figure 3B.11 Snapshot of a 10ns molecular dynamics simulation of compound 3.9 (blue) seated within the minor groove of the DRH structure. The molecule does not intercalate within the DRH structure, and fails to form any significant interactions with the DRH structure 185

Figure 3B.12 Image of the highest score (GBSA) docked pose of compound 3.5 (black) interacting with the F21T quadruplex (PDB ID: 3CDM) (blue and white). The benzimidazole moieties interact with the top and bottom faces of the quadruplex through van der Waals interactions, thus providing stabilisation 187

Figure 3B.13 Highest score (GBSA) docked pose of compound 3.9 (black) interacting with the F21T quadruplex (PDB ID: 3CDM) (blue and 188

white). The molecule does not form stacking interactions with the quadruplex due to a poor shape-fit

Figure 3B.14 Highest score (GBSA) docked pose of compound 3.39 (cyan) interacting with the F21T quadruplex (PDB ID: 3CDM) (blue and white). One fluoro-tetrahydroacridine moiety of the dimer forms stacking interactions with the DNA bases (blue nucleotides), whilst the second moiety fails to interact 189

Figure 3B.15 Hairpin structures of (a) the telomeric DNA/RNA hybrid duplex used in the CD studies (red indicates the RNA nucleotides) and (b) the control cDRH sequence 192

Figure 3B.16 CD spectrum of (a) the telomeric DNA/RNA hybrid duplex sequence (DRH), and (b) the control DNA duplex (cDRH) 192

Figure 3B.17 CD spectrum of compound 3.5 with (a) DRH (5 μM) and (b) cDRH in Tris buffer (pH 7.4) at 0-5 equivalent ligand concentration (path length 10 mm) 193

Figure 3B.18 CD spectrum of compound 3.6 with (a) DRH (5 μM) and (b) cDRH in Tris buffer (pH 7.4) at 0-5 equivalent ligand concentration (path length 10 mm) 194

Figure 3B.19 CD spectrum of compound 3.8 with (a) DRH (5 μM) and (b) cDRH in Tris buffer (pH 7.4) at 0-5 equivalent ligand concentration (path length 10 mm) 195

Figure 3B.20 CD spectrum of compound 3.7 with (a) DRH (5 μM) and (b) cDRH in Tris buffer (pH 7.4) at 0-5 equivalent ligand concentration (path length 10 mm) 195

Figure 3B.21 CD spectra of compounds (a) 3.4, (b) 3.9 and (c) 3.10 interacting with the DNA/RNA hybrid duplex sequence (5 μM) in Tris buffer (pH 7.4) at 0-5 equivalent ligand concentration (path length 10 mm) 196

Figure 3B.22 Highlighted features of the *bis*-benzimidazole molecules showing compound 3.6 as an example. The phenylene and biphenylene spacers are present in 3.9 and 3.10, respectively) 197

Figure 3B.23 CD spectrum of compound 3.40 with (a) DRH (5 μM) and (b) cDRH in Tris buffer (pH 7.4) at 0-5 equivalent ligand concentration (path length 10 mm) 198

Figure 3B.24 CD spectrum of compounds (a) 3.14, (b) 3.23, (c) 3.28 and (d) 3.35 intercalating with DRH (5 μM) in Tris buffer (pH 7.4) at 0-5 equivalent ligand concentration (path length 10 mm) 199

Figure 3B.25 CD spectrum of the interaction of compounds (a) 3.9 and (b) 3.10 with DRH (5 μM) in Tris buffer (pH 7.4) at 0-5 equivalent ligand concentration (path length 10 mm) 200

Figure 3B.26 CD spectrum of (a) human telomeric sequence ((d[$AG_3(TTAGGG)_3$])) in 50 mM Na-buffer and (b) telomeric G-quadruplex sequence (F21T) in Tris buffer (pH 7.4) in the presence of 100 mM NaCl 201

Figure 3B.27 CD spectrum of compounds (a) 3.4, (b) 3.5, (c) 3.6, (d) 3.7, (e) 3.8, (f) 3.9 and (g) 3.10 interacting with the F21T sequence in Tris buffer (pH 7.39) in the presence of 100 mM NaCl at 0-5 equivalents ligand concentration (path length 10 mm) 203

Figure 3B.28 CD spectrum of the interaction of compounds (a) 3.23 and (b) 3.69 with the telomeric G-quadruplex sequence (F21T) in Tris buffer (pH 7.39) in the presence of 100 mM NaCl at 0-5 equivalent ligand concentration (path length 10 mm) 204

Figure 3B.29 Reversed Transcriptase-PCR results for modulation of hTERT gene expression by (a) Compound 3.6 (1 μM), (b) Compound 3.8 (1 μM), (c) Compound 3.8 (5 μM), (d) Compound 3.43 (1 μM) and (e) Compound 3.71 (10 μM). La: 100 bp DNA ladder, C : cell control, L: Cell stimulated with LPS, D: Cell stimulated with LPS and compound, + : RT enzyme, -: No RT enzyme 213

Figure 3B.30 RT-PCR results for evaluation of the effect on Cyclin B gene expression in MDA-MB-231 cells by (a) Compound 3.6 (1 μM), (b) Compound 3.8 (1 μM), (c) Compound 3.8 (5 μM), (d) Compound 3.43 (1 μM) and (e) Compound 3.71 (10 μM). La: 100 bp DNA ladder, C: cell control, L: Cell stimulated with LPS, D: Cell stimulated with LPS and compound, +: RT enzyme, -: No RT enzyme 214

Figure 3B.31 Electrophoretic mobility shift assay using 4% agarose gel. Control Oligo-DRH (lane 1), Oligo-DRH with compound 3.6 (2.5 μM) (in lane 2), Oligo-DRH with compound 3.8 (2.5 μM) (in lane 3) 216

Appendix 2: Experimental data and Graphs

A2.1: FRET melting result 265

A2.2: UV calibration curves obtained on Circular Dichroism (CD) Studies 272

A2.3: IC_{50} curves observed in Cytotoxicity Studies (MTT Assay) 274

List of Tables

Chapter 1		
Table 1.1	Geometric comparison among A-DNA, B-DNA and Z-DNA	45
Table 1.2	Examples of DNA sequences identified as forming quadruplex structures in the human genome	69
Chapter 3A		
Table 3A.1	*Bis*-benzimidazole molecules in library-1	142
Table 3A.2	*Bis*-quinoline molecule in library-2A	147
Table 3A.3	*Bis*-(6-fluoro-2-methylquinoline) molecules in library-2B	151
Table 3A.4	*Bis*-indole molecules in library-2C	154
Table 3A.5	*Bis*-acridine type molecules in library-3A	157
Table 3A.6	*Bis*- (2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*] quinoline-11-carboxamide molecule in library-3B	160
Table 3A.7	*Bis*-benzofuran type molecules in library-4	166
Chapter 3B		
Table 3B.1	FRET melting assay results for library-1 molecules with telomeric DNA/RNA hybrid duplex (DRH) and control cDRH sequences	171
Table 3B.2	FRET melting assay results for library-1 molecules with telomeric G-quadruplex (F21T) and control duplex DNA (cDDS) sequences	172
Table 3B.3	FRET melting assay results for library-3B molecules with telomeric DNA/RNA hybrid duplex (DRH) and control cDRH sequences	173
Table 3B.4	FRET melting assay results for library-3B molecules with telomeric G-quadruplex (F21T) and control duplex DNA (cDDS) sequences	174
Table 3B.5	Free energy of binding calculations (kcal/mol) of the intercalation of library-1 molecules and NSC 273829 with DRH and cDRH in explicit solvent	181
Table 3B.6	Docking scores (GBSA) of selected members from libraries 1-3 interacting with the F21T quadruplex (PDB ID: 3CDM)	190
Table 3B.7	Cytotoxic activity of compounds from library-1, library-3B and library-4 in the MDA-MB-231 cell line	206
Table 3B.8	Cytotoxic activity of compounds from library-1, library-3B and library-4 in the HeLa cell line	208
Table 3B.9	Cytotoxic activity of compounds from library-1, library-3B and library-4 in the NCI H1975 cell line	209

Abbreviations

Abl1	Abelson murine leukaemia viral oncogene homolog 1
ALT	Alternative lengthening of telomeres
APB	ALT-associated promyelocytic leukaemia bodies
A-T	Adenine-Thymine
ATP	Adenosine triphosphate
ATRX	Alpha thalassemia/mental retardation syndrome X-linked (gene)
Bcl-2	B-cell lymphoma 2
BCR	Breakpoint cluster region
cDNA	Complementary DNA
cDRH	Control DNA duplex sequence, equivalent to DRH sequence (where RNA bases of DRH were replaced by corresponding DNA bases)
C-G	Cytosine-Guanine
CSF	(Macrophage) colony-stimulating factor-1
CXCF12	C-X-C motif chemokine 12 (The stromal cell-derived factor 1)
DAXX	Death domain-associated protein
DCM	Dichloromethane
DIC	1,3-Diisopropylcarbodiimide or *N,N'*-diisopropylcarbodiimide
DMF	*N,N*-Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DDR	Discoidin domain receptor (family)
DRH	Telomeric DNA/RNA hybrid duplex sequence
ECTRs	Extrachromosomal telomere repeats
EGF	Epidermal growth factor
EMSA	Electrophoretic mobility shift assay
ETS	E-twenty-six transformation-specific transcription factor
EtOAc	Ethyl acetate
F21T	Human Telomeric G-quadruplex sequence
FGF	(Basic) fibroblast growth factor
FRET	Fluorescence Resonance Energy Transfer
GMP	Guanosine monophosphate
HGF	Hepatocyte growth factor
HOBt	1-Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
hPot1	Human Pot1
hRap1	Human Ras-proximate-1 or Ras-related protein 1
HRMS	High Resolution Mass Spectrometry
hTelo	Human Telomeric DNA sequence
IL-1β	Interleukin-1 beta
IR	Infrared

LCMS	Liquid Chromatography-Mass Spectrometry
MAPK	Mitogen-Activated Protein Kinase
MeOH	Methanol
mRNA	Messenger RNA
MS	Mass Spectrometry
NAD(H)	Nicotinamide Adenine Dinucleotide (Hydride)
NADPH	Nicotinamide Adenine Dinucleotide Phosphate (Hydride)
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells, a protein complex
NMR	Nuclear Magnetic Resonance
NSE	Nuclease-Sensitive Element
PAX8	Paired box gene 8
Pd	Palladium
PDGF	Platelet-derived growth factor
PKB	Protein Kinase B, (alias Akt), is a serine/threonine-specific protein kinase
PNA	Peptide nucleic acid
POT1	Protection of Telomeres 1
pRb	Retinoblastoma protein
PTEN	Phosphatase and tensin homolog
Rb	Retinoblastoma
RNA	Ribonucleic Acid
RNP	Ribonucleoprotein
RPA	Replication Protein A
RT	Room temperature
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SC	Synaptonemal Complexes
TBP	TATA Binding Protein
TEBP	Telomere End-Binding Proteins
TER	Telomerase RNA
TERT	Telomerase Reverse Transcriptase
TGF-β	Transforming growth factor beta
TIN2	TRF1-interacting nuclear factor 2
TLC	Thin Layer Chromatography
Topo I	Human Topoisomerase I
TRAP	Telomerase Repeat Amplification Protocol Assay
TRF1	Telomeric Repeat binding Factor 1
TRF2	Telomeric Repeat binding Factor 2
TTTA Box	Goldstein-Hogness Box
UV	Ultraviolet
VEGF	Vascular endothelial growth factor
WT1	Wilms tumour protein

Abstract

Telomeres are repetitive sequences of DNA at the ends of chromosomes that become progressively shorter during cell division, acting as a form of “biological clock” causing cell death once they have reached a certain length. Almost 90% of cancer cells overexpress the enzyme telomerase which can lengthen telomeres and confer immortality to the tumour cells. Thus, telomerase has become an important target for drug discovery in the oncology area, and there is also interest from researchers investigating the aging process.

During the catalytic cycle of telomerase, a unique DNA/RNA hybrid duplex (DRH) forms that is typically between 6-11 base pairs long and is key to extending the telomere. There is interest in discovering small drug-like molecules that can recognize and bind to this hybrid duplex to inhibit selectively telomerase, either by stabilizing the structure and thereby preventing telomerase dissociation (a key step in the catalytic cycle) or by sufficiently distorting the hybrid duplex to cause the misalignment of key catalytic groups.

This project began by using oligonucleotides representing DNA/RNA hybrid duplex (DRH), telomeric G-quadruplex and control duplex DNA sequences to screen against the National Cancer Institute compound libraries (*i.e.*, Diversity Set II, Mechanistic Set and Natural Product Set) using a high throughput Fluorescent Resonance Energy Transfer (FRET)-based DNA thermal denaturation assay to determine binding affinity and specificity. Thirteen novel chemical scaffold families were identified in the assay, compounds which showed a >5 °C selective stabilization of the DNA/RNA hybrid duplex at a 1 μM ligand concentration. Chemical modifications were then made to these scaffolds to generate focused libraries of analogues to improve selectivity, potency and drug-likeness, and to provide Structure-Activity Relationship (SAR) information.

A total of 49 novel molecules were synthesized and then screened against an expanded range of four different nucleic acid constructs including telomeric and DNA/RNA hybrid duplex sequences. A number of compounds showed selective DNA/RNA hybrid stabilization potential with some compounds also showing notable telomeric G-quadruplex stabilization without significant affinity for promoter G-quadruplexes (*i.e.*, c-Kit1, c-Kit-2 and c-Myc) and control

duplex DNA sequences. The compounds from library-1 provided DNA/RNA hybrid duplex stabilization in the 0.5-7.2 °C range and telomeric G-quadruplex stabilization in the 0.2-6.5 °C range at a 1 μM ligand concentration. Molecular modelling and molecular dynamics studies confirmed that the methylene spacer between the benzimidazole and phenylene moieties of molecules within library-1 is perfectly shaped to fit within the DRH sequence. In addition, it was confirmed that minor-groove binding and simultaneous intercalation between the nucleobases of a DNA/RNA hybrid duplex requires a linker of specific length (*i.e.*, an eight methylene spacer as in compound **3.3**). Selected compounds were then studied further using a variety of biological techniques to confirm selective telomerase inhibition and cell-based assays to utilize their potential as antitumour agents.

Chapter 1: Introduction

1.1 Cancer and General Overview

In general, cancer is a condition where cells can grow and reproduce uncontrollably in a specific part (*i.e.*, cancer proliferation site) of the body. These cancerous cells can invade and destroy nearby healthy tissues and organs. Cancer cells usually begin in one part of the body and can be implanted into a distant site by metastasis. The metastasis stage only arises when malignant cells enter either the bloodstream or lymph system.

Cancer propagation comprises continuous changing and modifications to the cell genome carried out by both internal and external factors. Tumourigenesis is the process that leads to cancer cell formation, and an understanding of this process leads to the discovery of mutated genes (*i.e.*, oncogenes) and the loss of function of tumour suppressor genes. Tumourigenesis involves a multistep process where genetic changes happen at each step that ultimately lead to the progressive transformation of healthy cells into tumour cells. Studies confirm that such genetic modifications of tumour cells occur at different sites, ranging from point mutations (*i.e.,* one DNA base-pair change) to chromosomal translocations.

This uncontrolled cellular proliferation results in a preliminary heterogenic tumour, and the tumour cells ultimately undergo metaplasia, subsequent dysplasia and anaplasia that finally results in a malignant phenotype. Such malignancy allows for incursion into the blood circulation (metastasis), followed by invasion of a subsequent site for tumourigenesis. As soon as the tumour cells come to another resting site, they penetrate the new blood vessel walls to undergo further multiplication and ultimately form a clinically detectable tumour. This new tumour is known as a metastatic tumour.[1]

In 2012, Schroeder and co-workers reported metastasis as a cause of a vast majority of cancer related death.[2] They described the metastasis process as involving seven steps. At the beginning metastatic cells release from the primary tumour (**a**) and decrease the adhesion to adjoining cells, then clear a path for migration into the vasculature-rich stroma (**b**) and up on vasculature formation, cells can freely enter the bloodstream. Intravasation (**c**) is necessary

for continuous vasculature formation, where cells cause either endothelial cell retraction or endothelial cell death. In the bloodstream, cancer cell distribution happens due to blood flow and interactions between cancer cells and the secondary organs that they colonize (*i.e.*, cells can get trapped in narrow capillary beds, such as capillaries in the lung and liver), and can also express receptors that bind to metastasis-supporting sites (**d**) or to platelets (**e**), which help the cancer cells to escape from immune response. After reaching the secondary site, cancer cells can exit the bloodstream (**f**) by inducing endothelial cell retraction or death. To proliferate in the secondary site, cancer cells co-opt the local environment by releasing pro-inflammatory compounds and proteinases that induce their neighbours to release growth factors (**g**).

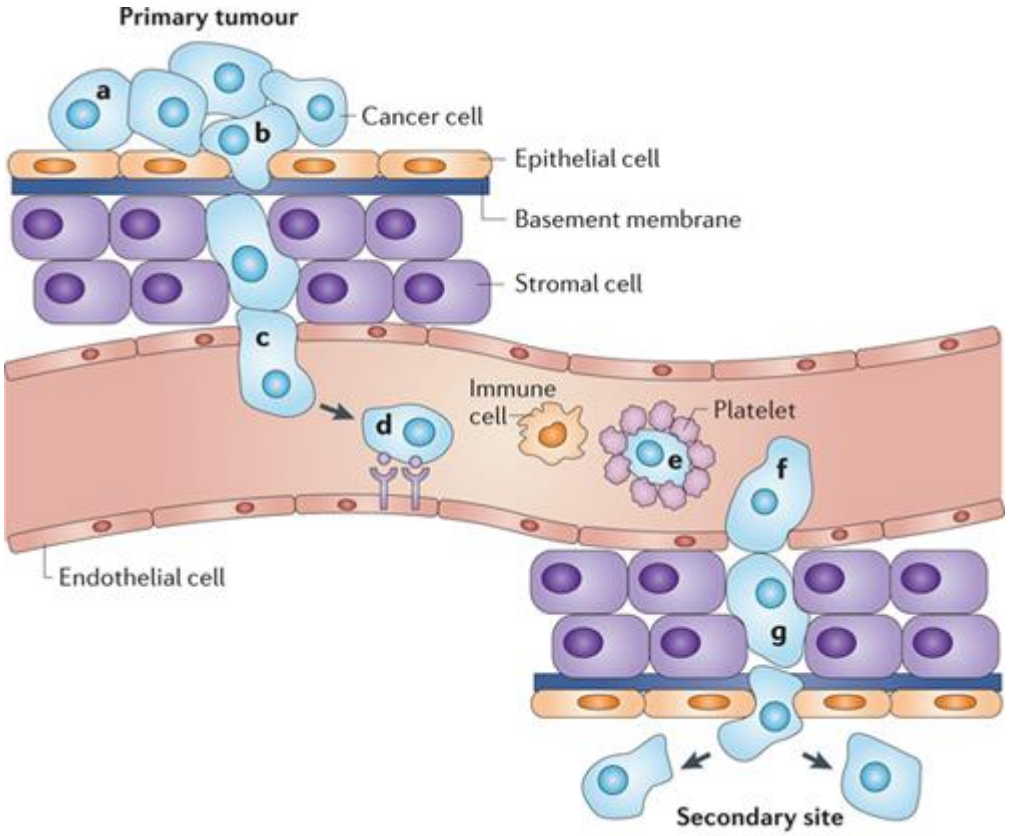


Figure 1.1: Cancer cells dispersion mechanism (*i.e.*, metastasis).

There are around 100 different types of cancer, which vary according to the origin of the cancer cells, their growth, the stages of disease propagation and the age of victims, and each of them has its own methods for diagnosis and treatment. A most noteworthy point is that not all tumours can undergo malignant transformation; some remain at the benign stage and do not grow at all.[3]

There are four main types of cancer[4]:

a. cancers of the organs (Carcinoma),
b. cancers of the muscles, bone, cartilage, and connective tissue (Sarcoma),
c. cancers of the lymphatic system (Lymphoma), and
d. cancers of the blood-making system (Leukaemia)

It is now widely believed that cancer is a result of the accumulation of mutations in the genes that directly control cell birth or cell death. The mechanisms through which these genetic transformations originate are still the subject of ongoing debate. It is generally said that an underlying genetic instability is the key starting point of the multiple mutations and subsequent cancer propagation.[5, 6]

Cancer cells vary according to their patterns of growth and spread throughout the body. The American Joint Committee on Cancer defined the stages of cancer growth as follows:

1. Stage I – Primary tumour only.
2. Stage II – Primary tumour, but larger than in Stage I.
3. Stage III – Primary tumour and metastasis to lymph nodes.
4. Stage IV – Primary tumour and distant metastasis.

According to the World Cancer Report 2014, cancers are the leading causes of morbidity and mortality worldwide, where tobacco is held to cause approximately 70% of global lung cancer related deaths and 20% of overall cancer deaths. The WHO also reported that lung, prostate, colorectal, stomach and liver cancers are predominant in men and breast, colorectal, lung, cervix and stomach cancers are predominant in women. High BMI, low fruit and vegetable consumption, lack of physical movement and use of tobacco and alcohol were all highlighted in the report as leading behavioural and dietary risk factors for one-third of cancer deaths.[7]

1.2 Hallmarks of Cancer

According to Hanahan and co-workers, all types of cancer exhibit six common traits that direct the transformation of normal cells to cancer cells.[8] These traits are well-known as hallmarks of cancer, with the six biological capabilities attained during the multistep progression of human tumours. The six traits include sustained proliferative signalling, escaping growth suppressors, resisting cell death, facilitated replicative immortality, induced angiogenesis and activated incursion and metastasis. Two more promising hallmarks were listed by Hanahan and colleagues in 2011[1]; reprogramming of energy metabolism and escaping immune destruction. They also mentioned genome stability and mutation and tumour promoting inflammation as two enabling traits for cancer.

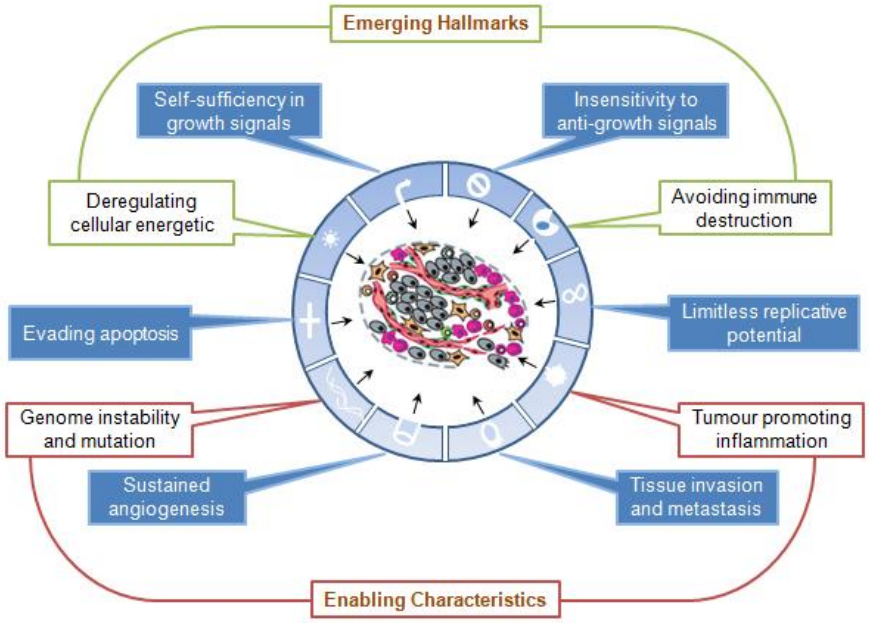


Figure 1.2: Reproduced illustration of the six hallmarks of cancer and two emerging hallmarks of cancer, adapted from Hanahan and co-workers.

1.2.1 Sustaining proliferative signalling

Cancer cells acquire self-sufficiency in growth signals to reduce their dependence on external factors (*e.g.*, biochemical stimulation from cell environment like hormones and other molecules). Cancer cells gain some capabilities in a multiple way such as by generating autocrine signalling (*i.e.*, by permanently activating the signalling pathways that

respond to these signals), by destroying ‘off switches' to prevent excessive growth (*i.e.*, negative feedback). Such deregulation in cell division occurs due to the alteration of the responsible protein.

Defects in negative-feedback mechanisms increase proliferative signalling. Usually, this type of regulation involves Ras-oncoproteins and oncogenic mutations affecting Ras-genes, which compromise Ras GTPase activity. Related negative-feedback mechanisms also function at multiple nodes inside the proliferative signalling circuitry like PTEN phosphatase, where mutations in PTEN intensify the PI3-Kinase signal. In this way, such mutation can promote tumourigenesis, which has been evaluated in a variety of experimental models of cancer.[9]

1.2.2 Evading growth suppressors

Typically, cancer cells are resistant to growth-preventing signals from their neighbours. Tumour suppressor genes are responsible for controlling cell division, as they carry a signal from the cell to confirm the cell division necessity, and also have capability to halt cell division according to their need (*e.g.*, in case of damage DNA). However, in cancer, due to the alteration of tumour suppressor proteins, they lose the successive capability to prevent cell division, even when the cell has severe abnormalities. Moreover, cancer cells usually lose the contact inhibition property (*i.e.*, another way to prevent excess cell division in normal cells, where cells stop dividing as soon as they fill up the space within a certain area and touch each other), which leads to continuous grow and cell division independent of their neighbourhood. Some cancer cells also escape the quiescent state (G0) of the cell cycle, where cells become insensitive to extracellular antigrowth signals.

Transforming growth factor beta (TGF- β) is also well known for its anti-proliferative activity, but TGF-β signalling pathways are redirected away from restraining cell proliferation in many late-stage tumours, causing activation of the cellular process known as epithelial-to-mesenchymal transition that is associated with high-grade malignancy.[10]

1.2.3 Evading programmed cell death or apoptosis

Apoptosis is a mechanism by which cells are programmed to die in the event of cell damage, whereas cancer cells have the ability to bypass this mechanism. Moreover, apoptosis is a

requirement for organisms to grow and develop properly and to maintain body tissues. It is assumed that the cancer cells may acquire some abilities by altering the mechanisms that detect the damage or abnormalities within cells. Therefore, evasion of apoptosis is another trait for a tumour to expand.

1.2.4 Limitless replicative potential

Non-cancer cells die after a certain number of divisions, which is an intrinsic program, the Hayflick limit (*i.e.*, that limits cells' multiplication to about 60-70 doublings, at which point cells reach a stage of senescence). On the other hand, cancer cells can escape this limit and are capable of indefinite growth and division that lead to immortality. Studies confirmed that such immortal cells possess damaged chromosomes, primarily at their ends (*i.e.*, telomeres). Typically, telomeric DNA shortens with every cell division, until it becomes short enough to activate senescence. Cancer cells can overcome this limit by restricting pRB and p53 tumour suppressor proteins. Recent studies also confirmed that many cancers (*i.e.*, approximately 90%) are involved with the upregulation of telomerase enzyme, which is responsible for maintaining telomeres length elongation by using the RNA template within the enzyme itself.

1.2.5 Sustained angiogenesis

Angiogenesis is the process of forming new blood vessels. Cancer cells are able to carry out this process, to ensure a continuous supply of oxygen and other nutrients as well as removal of waste products. Cancer cells acquire the capacity to produce new vasculature by activating the 'angiogenic switch' by controlling non-cancerous cells present in the tumour by reducing the production of factors that inhibit blood vessel production, and increasing the production of factors that promote blood vessel formation.

1.2.6 Tissue invasion and metastasis

Cancer cells can disrupt away from their site of origin and can invade surrounding tissue, which subsequently can spread to distant body parts. Such metastases processes lead to an

untreatable stage for cancer treatment strategies. Approximately 90% of cancer deaths are caused by metastasized tumour growth and consecutive diseases.

In 2011, Hanahan and co-workers postulated further four traits as hallmarks of cancer, which were classified as 'enabling' and 'emerging' characteristics for cancer. The four new traits are as follows:

a. ***Deregulation of cellular energetics***: Most cancer cells utilize atypical metabolic pathways to generate excess energy for fast growing cancer cells. Cancer cells gain the ability to reprogram their energy production from the normal aerobic metabolism of glucose to glycolysis. Additionally, a variety of generated intermediates from glycolysis can assist in the biosynthesis of tumour-promoting macromolecules.

b. ***Evading immune system***: Cancer cells can become invisible to the body defence mechanism (*i.e.*, immune system) to bypass destructive consequences. Hanahan suggested this as a new evolving characteristic for cancer.

c. ***Genome instability and mutation***: In general, cancer cells acquire abnormalities in chromosomal structure. Such genomic instabilities and mutations can occur due to an increase in the rate of mutation caused by sensitivity to mutagenic agents or reduced efficiency of genome maintenance systems which includes DNA repair enzymes. Somatic mutations can activate downstream pathways in the cancer cell cycle. High-throughput DNA sequencing studies of cancer cell genomes have shown that constitutive initiation of signalling circuits is triggered by the activated growth factor receptors due to somatic mutations in human tumours. It is also reported that approximately 40% of human malignancies contain stimulated mutations that influence the B-Raf protein structure, which causes simultaneous signalling through the Raf-protein to mitogen activated protein kinase (MAPK).[11] Moreover, mutations in the catalytic subunit of phosphoinositide 3-kinase (PI3-kinase) hyper-activate the PI3-kinase signalling pathway and Akt/PKB signal, which were also observed in a group of tumours.[9] For this reason, genomic instability and mutation are considered as enabling characteristics of cancer.

d. ***Tumour-promoting inflammation***: Recent studies demonstrated that the inflammation process can promote tumourigenesis by providing growth factors and thus facilitate the angiogenesis process.

1.3 Treatment Options for Cancer

There are number of treatment options available for cancer, where the main approach of any treatment is to remove and/or kill cancerous cells to achieve non-recurrence of cancer cell development. However, it is challenging to ensure either the complete removal or cure of cancer cells as cancer cell proliferation is caused by various cell signalling pathways. Thus, researchers are investing their knowledge to develop a single treatment approach to evade cancer. To date, the main treatments for cancer are surgery, radiotherapy and small-molecule chemotherapy.

1.3.1 Chemotherapy

From the beginning, chemotherapy was accepted as one strategy in combating cancer cells and tissues. The fundamental principle of this approach is to kill cancer cells by treating them with chemicals that can obstruct susceptible processes related to cell division. Medicines (chemotherapeutic agents) treat cancer cells and/or suppress cancer cell proliferation either by destroying important proteins involved in cell signalling pathways or by damaging the DNA itself.[12, 13] The nitrogen mustards were the first chemical agents with low molecular weight, which were introduced clinically in the 1940s as chemotherapeutic agents with anti-leukemic properties. Promising developments have been observed in the discovery of new anticancer drugs during the last few decades when methotrexate and cisplatin came upfront as chemotherapeutic agents. However, while chemotherapy drugs are effective against cancer cells, they simultaneously affect rapidly proliferating normal cells such as those connected to the bone marrow, immune system, gastro-intestinal tract and hair follicles. As a result, chemotherapy consistently brings a number of unwanted side effects, for example biliousness, alopecia, pain and vomiting. For these shortcomings, researchers are trying to develop small molecules that can act as site-specific drugs and interrupt DNA replication, cell division and tumour growth in cancer cell-oriented locations.

Clinically used chemotherapeutic agents have a wide variety of mechanisms in order to interfere with a number of cancer development processes such as tumour cell growth, motility, survival, and angiogenesis. However, the DNA interfering small molecules possess the largest and most varied ways of anticancer functionalities. There are also other mechanisms involved in anticancer treatment strategy.

1.3.1.1 DNA targeted agents

Cancer causes different changes to the structure of DNA, causing alteration of normal DNA regulation. DNA has been targeted since the invention of chemotherapeutic agents and these drugs primarily work by binding to DNA and subsequently inhibiting gene transcription, translation and other cellular processes.

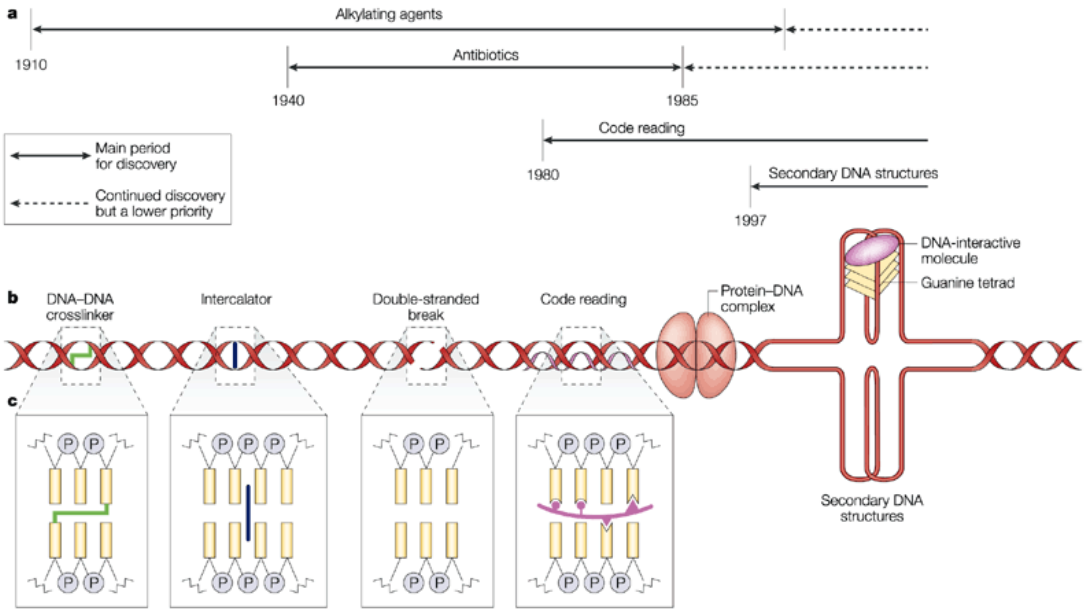


Figure 1.3: (**a**) A timeline for the discovery and evolution of DNA-interactive agents. (**b**) Illustrates the types of modifications that can form on DNA (from left to right, crosslinks, intercalation, DNA-strand cleavage and code-reading molecules). The DNA helix is also associated with proteins (e.g. topoisomerase II) and secondary DNA structures (e.g. G-quadruplex). (**c**) Detailed illustration of the different types of DNA modification. (Adapted from DNA and its associated processes as targets for cancer therapy by Laurence H. Hurley)[14]

A large proportion of chemotherapeutic anticancer drugs can interact with DNA either directly or can prevent the appropriate relaxation of DNA by inhibiting topoisomerases (enzymes that regulate the over-winding or under-winding of the DNA helix).

In cancer treatment, Drug/DNA interactions can be accomplished in two major ways, *intercalation* and/or *groove binding*. Intercalation involves the placing or stacking of a planar molecule between DNA base pairs and causes a decrease in the helical twist that leads to broadening of the DNA and, consequently, alteration of the regular helical DNA structure.[15] There is an important free energy involved in making this intercalation cavity, which is approximately 4 $kcalmol^{-1}$ and positive contributions of hydrophobic interactions, ionic bonding, hydrogen bonding and van der Waal's interactions result in association constants of 10^5-$10^{11}M^{-1}$.[16] Moreover, the intercalation process is usually associated with molecules containing fused bi- and/or tri-cyclic ring systems. According to Chaires and co-workers, binding does not cause big conformational changes in DNA and can be described according to typical lock-and-key models for ligand/macromolecular binding.[16]

In contrast, groove binding agents are usually semi-circular shaped molecules that can fit and bind to the minor groove of DNA helices (**Figure 1.15**). Sometimes, such small molecules possess both modes of binding efficacy due to them having an intercalating unit along with groove-binding side arms in the same molecule.

A large number of chemotherapeutic agents exert their effect by interacting with DNA. Typically, DNA targeted agents interfere with DNA processing (*i.e*., transcription, translation, replication), which lead to cell death through apoptosis. DNA interactive agents can be DNA alkylating agents (*e.g.*, dicarbazine, temozolomide, chlorambucil), cross-linkers with two DNA-alkylating moieties (*i.e.,* interstrand cross-linkers like nitrogen mustard, cisplatin and intrastrand cross-linkers like pyrrolobenzodiazepine dimer SJG-136), intercalators (*e.g.*, ethidium bromide, doxorubicin and other quadruplex intercalating agents), topoisomerase-I inhibitors (*e.g.*, belotecan, topotecan), topoisomerase-II inhibitors (*e.g.*, podophyllotoxin, ellipticine), and DNA cleaving agents (*e.g.*, bleomycin). It is to be mentioned that all the DNA interactive agents, except intercalators, interfere with the DNA structure either through the

minor or major groove of the DNA helix, thus in broad DNA interactive agents are also classified as DNA minor groove and major groove binding agents.

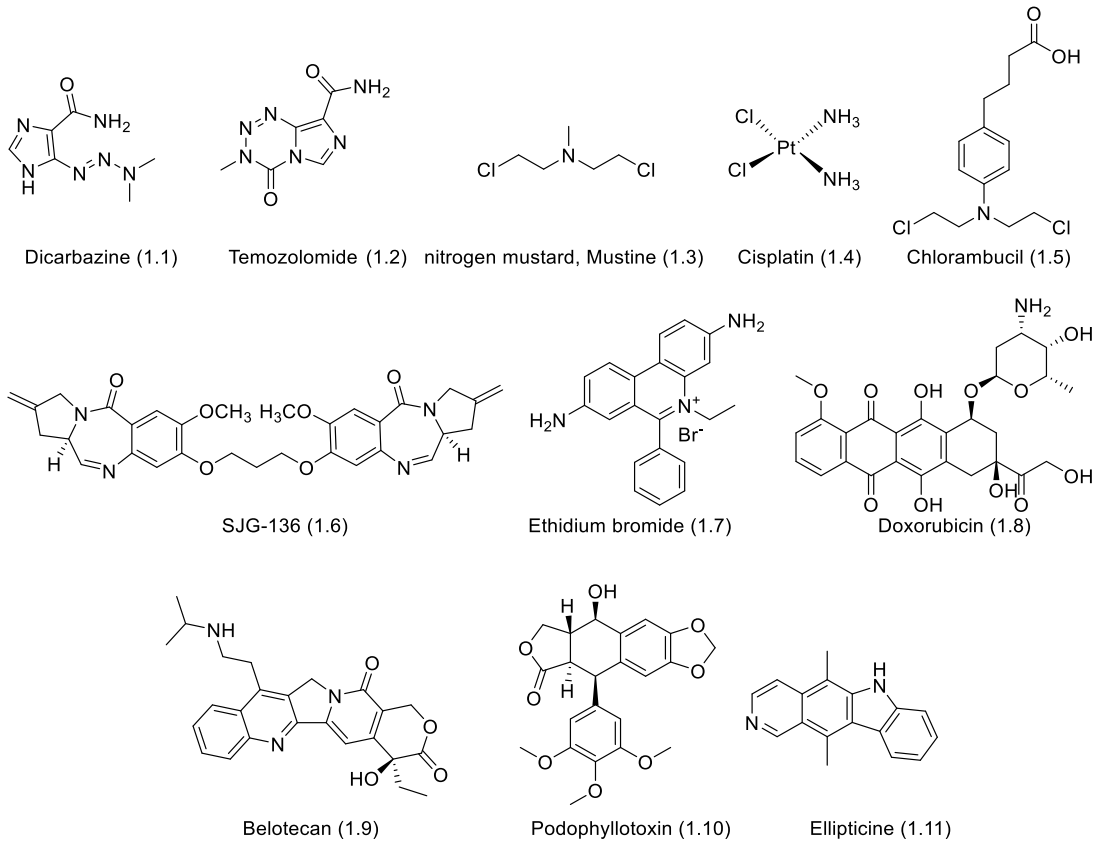


Figure 1.4: Structures of some DNA interacting agents.

Quarfloxin (also known as Itarnafloxin, **1.12**) is now in Phase-II clinical trials for the treatment of carcinoid and/or neuroendocrine tumours, and is recognised as a G-quadruplex stabilizing agent. It acts through inhibiting RNA polymerase (Pol-I)-based transcription by displacing nucleolin, a phosphoprotein involved in the synthesis and maturation of the ribosome, from the nucleus.[17]

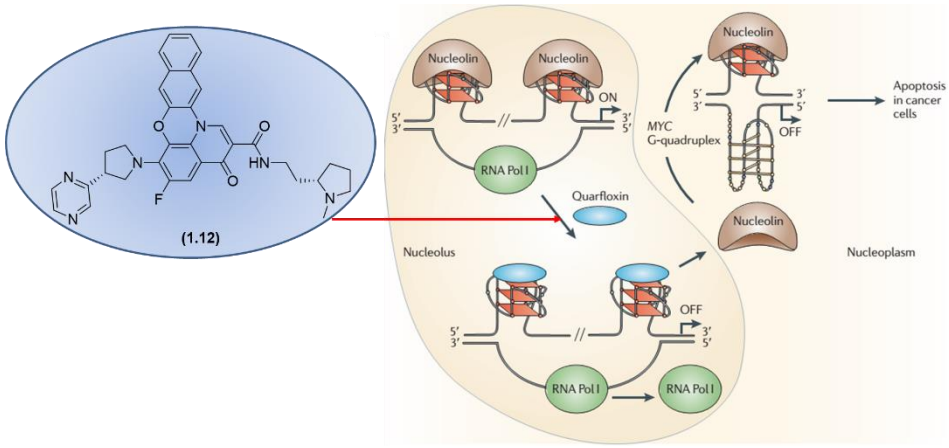


Figure 1.5: The structure of quarfloxin (**1.12**) and its mechanism of action.[18]

1.3.1.2 Antimetabolites

Antimetabolites are synthetic molecules that can block vital metabolic pathways which are required for cell growth. As tumour cells can proliferate at a faster rate than normal cell populations (exceptions are observed for bone marrow, hair follicle, GI tract), antimetabolites are believed to have antiproliferative properties. Methotrexate is one of the classical examples of this class of chemotherapeutic agent. Studies confirmed that the enzyme dihydrofolate reductase (DHFR) is responsible for the formation of tetrahydrofolic acid from dihydrofolic acid, which is vital for thymine synthesis. Methotrexate, a modified structure of dihydrofolic acid, competitively binds to the active site of DHFR and thus blocks the natural activity of dihydrofolic acid and subsequently leads to blockage of thymine synthesis.[19]

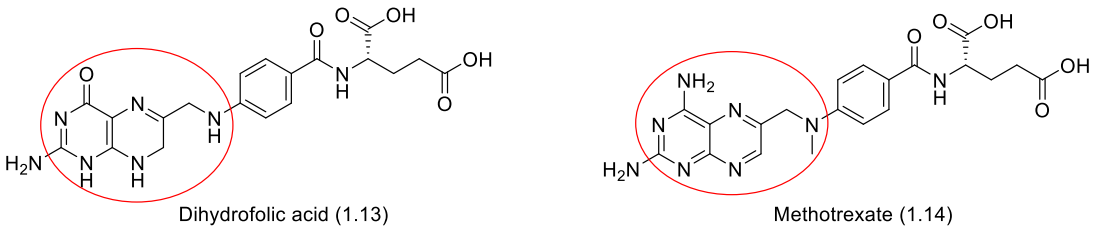
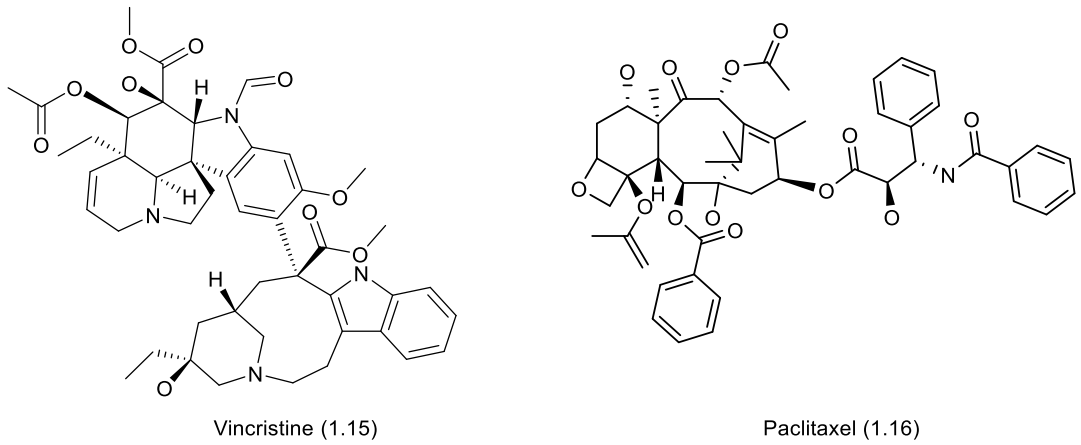


Figure 1.6: Structural differences between dihydrofolic acid and methotrexate (circled in red)

Other types of agents in this family include purine antimetabolites (*e.g.*, 6-Mercaptopurine), pyrimidine antimetabolites (*e.g.*, 5-fluorouracil[20]), thymidylate synthase inhibitor (*e.g.*, raltitrexed), adenosine deaminase (*e.g.*, pentostatin) and ribonucleotide reductase inhibitors (*e.g.*, hydroxycarbamide).[21]

1.3.1.3 Antitubulin Agents

Microtubules are filamentous intracellular assemblies which are involved in nucleic acid and cell division, organization of intracellular structure, and intracellular transport. Typically, microtubules are found as polymers of tubulin (*i.e.*, α-tubilin and β-tubilin) that are integral components of the mitotic spindle. Because the functions of microtubules are so critical to the existence of eukaryotic cells, agents that disturb this equilibrium (*i.e.*, assembly and disassembly) can block mitosis and lead to cell death.[22] The Vinca alkaloids (*e.g.*, vincristine and vinblastine extracted from *Vinca rosea*) and taxans (*e.g.*, paclitaxel[23]) are the most common example of this class of chemotherapeutic agents. Vinflunine, a novel fluorinated vinca alkaloid, is in Phase II clinical trials which blocks mitosis at the metaphase/anaphase transition and leads to apoptosis.[24]



Vincristine (1.15) Paclitaxel (1.16)

Figure 1.7: Structure of vincristine and paclitaxel.

1.3.1.4 Molecularly targeted agents

Generally, in cancer, molecularly targeted therapy means to a mode of medication that block the growth and spread of cancer by interfering with specific molecules (*i.e.*, enzyme, regulatory protein kinases, macromolecules within tumour cell) that are threatening for tumour development. Such a mechanism offers targeted cancer therapy to be equally or even more effective than chemotherapy and radiotherapy and less harmful to normal cells. This class of drugs are sometimes noted as signal transduction inhibitors or secondary messenger inhibitors. Research has shown that numerous cell signalling pathways are responsible for intra- and inter-cellular mechanisms like cellular growth, cell death (or apoptosis), and protein

synthesis, which are altered within cancer cells due to mutation. Thus inhibition of such signalling pathways can stop the formation of further mutated cells or tumourigenic cells, which can lead to an anticancer effect.[25] Molecularly targeted agents can be as follows according to either their targeted signal pathway or inhibition mechanism:

a. **Protein kinase inhibitors:** A protein kinase inhibitor is one type of enzyme inhibitor that blocks the action of one or more protein kinases, which responsible in the addition of a phosphate group to a protein and modulation of protein's function. Such phosphorylation regulates many biological processes, and thus protein kinase inhibitors are used to treat diseases related to hyperactivity of protein kinases (*e.g.*, in cancer). Few prominent examples are BCR-ABL inhibitor Imatinib, HER2/neu inhibitor Trastuzumab, EGFR inhibitor Erlotinib.[26]

b. **RAS pathway signal inhibitors:** The RAS proteins regulate the signalling pathways, responsible for some features of normal cell growth and malignant transformation. They are abnormal in most human tumours due to mutations in the *RAS* genes themselves or due to changes in upstream or downstream signalling mechanisms. Rational therapies targeting the RAS pathways can inhibit tumour growth and existence. Several of new therapeutic agents are showing potential in the chemotherapeutic drug development area, for example, Tipifarnib, BMS-214662, ARRY-142886).[27]

c. **Cell cycle inhibitors:** The cell cycle directs the transition from latency through cell growth to proliferation for each cell and the cyclin-dependent kinases (cdks) are the key regulators of the cell cycle machinery. In cancer therapy, the cdks can be the rational targets due to their expression in cancer cells is abnormal and their inhibition can tempt to cell death. Several drugs targeting the cell cycle have already in clinical trials. Some examples include flavopiridol, indisulam, AZD5438, SNS-032, bryostatin-1 and SCH 727965.[28]

d. **Proteosome inhibitors:** Proteasome inhibitors usually block the activity of proteasomes, which is responsible for breakdown of unwanted proteins. Though there are multiple mechanisms are involved, but inhibiting proteasome may lead to prevention of degradation of pro-apoptotic factors (*e.g.*, p53 protein, which is a

requirement for programmed cell death). Proteasome inhibitors are being considered in the treatment of cancer, and three are approved for use in multiple myeloma, such as bortezomib (Velcade™)[29], Carfilzomib (Kyprolis™) and Ixazomib (Ninlaro™).

To avoid undesirable side effects, molecularly targeted agent development and research is ongoing to find less toxic drug candidates.

1.3.1.5 Hormonal therapy

Hormones are signalling molecules produced by glands (comprised of the endocrine signalling system) in multicellular organisms which are transported to target distant organs by the circulatory system. Binding of the hormone to the receptor can result in the activation of a signal transduction pathway, which can result in regular cell functioning that includes quick non-genomic effects or slower genomic responses. Thus, hormones acting through their receptors result in increased specific protein synthesis through the activation of gene transcription. Hormone therapy can play a vital in the treatment of hormone dependent tissue growth in breast, endometrium, and testes.[30] Tamoxifen, a non-steroidal estrogen antagonist, is a licenced hormonal therapy used in the UK for breast cancer treatment.[31] Aromatase inhibitors (*e.g.*, aminoglutethimide, formestane, exemestane, letrozole), gonadorelin analogues (*e.g.*, buserelin, triptorelin, cetrorelix), anti-androgen agents (*e.g.*, cyproterone acetate, utamide, bicalutamide), and somatostatin analogues (*e.g.*, octreotide, lantreotide) are named treatment options belonging to this class of therapy. Not only antagonistic treatment but also agonistic treatment candidates are available for hormone therapy as an anticancer strategy, where diethylstilbestrol, gestonorone caproate, medroxyprogesterone acetate, and norethoisterone are the named treatment options.[32]

One of the greatest challenges in the development and discovery of chemotherapeutic drugs is the selectivity between tumour (cancer) cells and healthy tissues. Most of the clinically used chemotherapeutic agents are associated with a varied degree of side effects to the healthy tissues. To minimize the toxicity, researchers are making efforts to sequence the selective anticancer agents' design and development, wherein targeted DNA sequences are of the most

effective type in clinical use that could help in significant increases in the survival of cancer patients.

1.3.2 Antibody-drug conjugates (ADC) for cancer therapy

Monoclonal antibodies are used as an integral tool in research due to their high specificity and affinity for target antigens. Such monoclonal antibodies can impart therapeutic benefit by binding and inhibiting the function of target antigens: for example, trastuzumab (Herceptin®) is used as a chemotherapeutic agent. However, antibodies against tumour-specific antigens often lack therapeutic activity. Such drawbacks can be overcome by conjugating either a cytotoxic drug or radionuclide with the antibody and thus expand its utility by improving its potency and effectiveness. Such antibodies can then be used as a means to target and deliver a toxic payload to the selected diseased tissue. In the antibody-drug conjugates (ADC) system, monoclonal antibodies are conjugated to chemotherapeutic agents, where the antibodies used are specific to tumour cell-surface proteins. Thus this treatment is thought to be tumour specific. To design an effective ADC for cancer therapy, the pre-requisite requirements for selection of a proper target, are a monoclonal antibody against that target, potent cytotoxic molecules and conjugation of the monoclonal antibody to the cytotoxic agents. Considerable advances in all these features in the past few years have resulted in approval of ado-trastuzumab emtansine (Kadcyla®) and brentuximab vedotin (Adcetris®) for clinical use.[33]

1.3.3 Radiotherapy

High-energy particles or waves are used in radiotherapy, which is one of the most common treatments used to destroy cancer cells. Radiotherapy utilizes either X-rays or radiopharmaceuticals (*i.e.*, radionuclides) as sources of γ-rays. Theoretically, as tumour cells are hypoxic and are less sensitive to damage by irradiation (which works through formation of DNA-damaging oxygen free radicals), oxygen is administered to sensitize the tumour cells prior to (and during) radiation therapy, A highly focused beam of X-rays is applied locally to avoid damaging healthy tissue. In the case of γ-rays, cobalt-60, gold-198, and iodine-131 are used as radionuclides.

Sometimes, a chemo-radiotherapy approach is undertaken as a treatment combination of chemotherapy agents and radiation. For example, the antibody cetuximab (Erbitux™) is used in combination with radiotherapy for squamous cell cancer of the head and neck.

1.3.4 Biotherapy or immunotherapy

In treatment with immunotherapy, certain biological components (*e.g.*, antibody, antibody-drug conjugates, interferons, interleukins, enzymes, vaccines) are used to stimulate one's own immune system to attack cancer cells. Some types of immunotherapy are also sometimes denoted as biologic therapy or biotherapy. For example, specific tumour antigens are produced in breast cancer on their cell surfaces, and monoclonal antibodies specifically target these tumours (*e.g.,* Herceptin™) are being used in such types of cancer.[34]

1.3.5 Photodynamic therapy (PDT)

Photodynamic therapy (PDT) is a cancer treatment option where a photosensitizing agent (*e.g.*, Photofrin™) is used along with light to kill cancer cells. This photosensitizing agent only works upon activation in the presence of certain type of light (*i.e.*, light of a specific and appropriate wavelength). In principle, upon decay of the photosensitizing agent to its ground state, the available oxygen is transformed into the singlet form (*i.e.,* free radical), which is highly cytotoxic and damages the tumour cells.

1.3.6 Surgery

A surgical process is usually used to remove a tumour that is small in size and reasonably well located. In this process, a well-defined tumour is removed along with some adjacent tissues to confirm complete removal of the tumour. However, supplementary treatment by either chemotherapy or radiotherapy is provided to eliminate any remainder or metastasized cancer cells. Sometimes either radiotherapy or chemotherapy is applied prior to surgery to shrink the tumour to expedite its easy removal.

1.4 DNA Structure

Deoxyribonucleic acid (DNA) is a macromolecule that contains the genetic information used in the development, functioning and reproduction of all living organisms. DNA, together with proteins and carbohydrates, constitute the three major and vital macromolecules that are

essential for all forms of life. According to the Watson and Crick model, DNA molecules are polynucleotides, which are composed of simpler units (nucleotides). The two biopolymer strands are coiled around each other running in opposite directions. Each nucleotide is composed of nucleobases, either cytosine (C), guanine (G), adenine (A), or thymine (T), along with a deoxyribose sugar and a phosphate group. DNA is composed of four deoxyribonucleotides, that are classified as either purine (deoxyadenosine, dA and deoxyguanosine, dG) or pyrimidine (deoxythymidine, dT and deoxycytidine, dC) nucleotides, as well as the deoxyribose sugars and a phosphate backbone. The sugar carbons are denoted as C1′, C2′, C3′, C4′, C5′, starting with the carbon at the glycosidic bond that attaches the base to the sugar, to make them distinguishable from the carbons of the nucleobases (**Figure 1.8**). The deoxynucleotides of DNA lack an O2′ oxygen, which distinguishes them from ribonucleotides (RNA). The nucleotides are connected to one another in a sequential manner by covalent bonding between the sugar of one nucleotide and the phosphate of the following nucleotide and form an alternating sugar-phosphate backbone. It has been well-established that adenine binds with thymine, and cytosine binds with guanine through hydrogen bonding between the two separate polynucleotide strands that make up a double-stranded DNA helix.[35, 36]

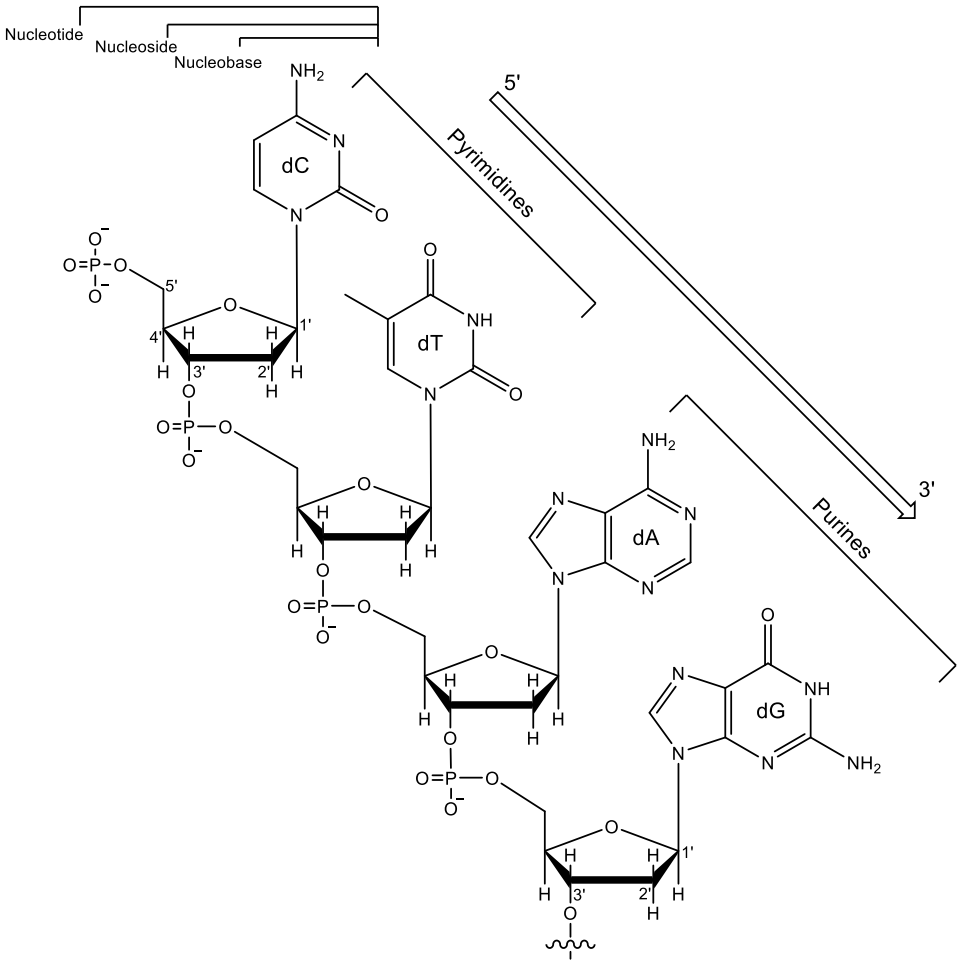


Figure 1.8: Components of DNA; the four common deoxyribonucleotides connected by phosphodiester bonds to form a single-strand (5′→3′).

In general, the DNA sequence is assembled through a condensation reaction to form a phosphodiester link that bridges the O3′ and O5′ oxygens of sequential nucleotides (**Figure 1.8**). It is common in the primary structure of a DNA strand that the link starts from the free O5′ oxygen (5′-end) and proceeds to the free O3′ oxygen (3′-end). In double stranded DNA, two corresponding strands are joined together in a sequence-specific style to form an antiparallel duplex sequence, where one strand is in the 5′ to 3′ direction and the counterpart 3′ to 5′.

DNA molecules are polymorphic and can adopt more than one form. About 60 years ago, x-ray diffraction studies carried out by Franklin and Wilkins discovered the A-form diffraction pattern produced by the dried DNA fibre; the same fibre produced a different, B-pattern when

moist. The B-form DNA structure is the familiar double helical DNA structure in which the axis of the helix runs through the Watson-Crick base pairs.[37]

Nucleic acid structures are also found to exist in triple-stranded and quadruple-stranded forms as well as left-handed forms. Quadruplex structures based on G-quartets are reported to form at the guanine-rich sequence of both the telomere and centromere regions whereas their complementary cytosine-rich strands are found to fold into an intercalated tetramer called the i-motif.[38]

Additionally, other DNA sequences have also been reported that can adopt a variety of unusual conformations; for example, polypurine-polypyrimidine rich sequences can form a triple-stranded structure known as H-DNA that is found abundantly in the eukaryotic genome, although their role has not been confirmed to date.[39]

1.4.1 Double Helical DNA

James D. Watson and Francis Crick first published the double-helix model of DNA structure (**Figure 1.9 a**) in 1953.[36] DNA, a long chain made from repeating units known as nucleotides, is 22–26 Å wide where each nucleotide is 3.3 Å long.[40] In a double helix, the strands are anti-parallel and the asymmetric ends of the strands are termed the 5′ end (with a terminal primary –OH group) and 3′ end (with a terminal secondary –OH group). The four nucleobases present in DNA, adenine (dA), cytosine (dC), guanine (dG) and thymine (dT), are joined to the sugar and phosphate and complete the nucleotide (**Figure 1.9 b**). The DNA double helix structure is stabilized through both hydrogen bonding between the nucleotides and base-stacking interactions between the aromatic nucleobases.[41]

The most accepted parameters for DNA duplex strands are the helical repeat (*i.e.*, the number of base pairs in one whole turn) and the helical rise (*i.e.*, the distance between nucleotides, measured along the helical axis). The repeat outlines the angle involving each base pair with the helical axis, where the helical twist is 360° per repeat that allows the repeat and rise (*i.e.*, the distance between one whole turn) of the DNA. Thus such parameters limit the DNA geometries. On the basis of varied geometric measures of the duplex DNA strand, three major types of duplex orientation are observed; A-DNA, B-DNA and Z-DNA.

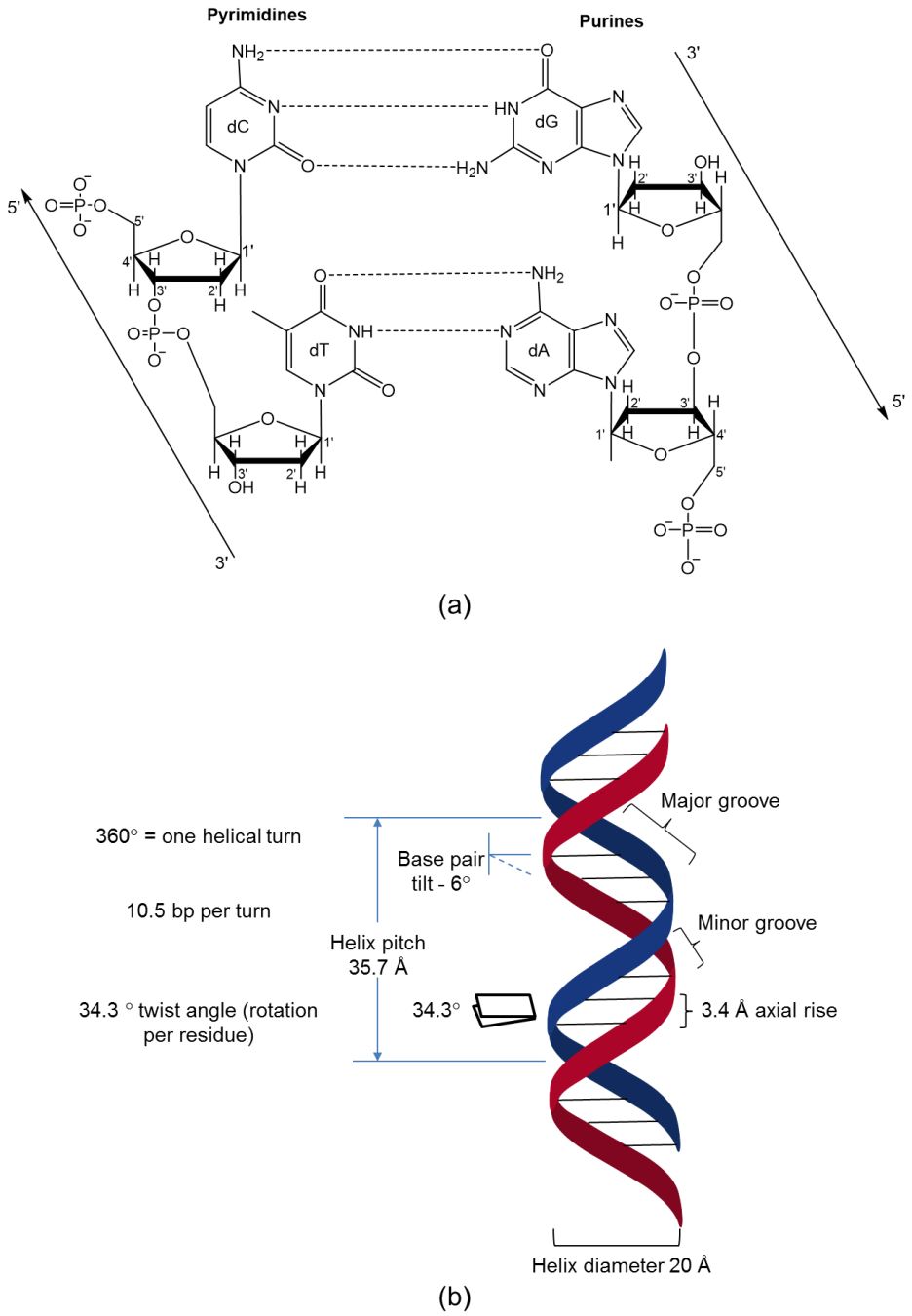


Figure 1.9: (a) Complementary base pairs in the DNA double helix, and (b) double stranded DNA helical structure.[39, 42]

1.4.1.1 A-DNA

Duplex DNA can adopt an A-form helical structure, a right-handed double helix similar to the more common form of DNA, B-DNA. It has been confirmed that if the deoxyribose sugar adopts a C3′-endo conformation with an approximate 6 Å phosphate-to-phosphate distance, this

results in a structure with a helical twist of around 31° that can accommodate a helical repeat of 11-12 base pairs, leading to a shorter, more compact helical structure known as A-DNA. Rosalind Franklin discovered the changes in conformation and also showed that under dehydrating conditions (*i.e.*, during crystal formation), regular DNA transforms into A-form DNA.

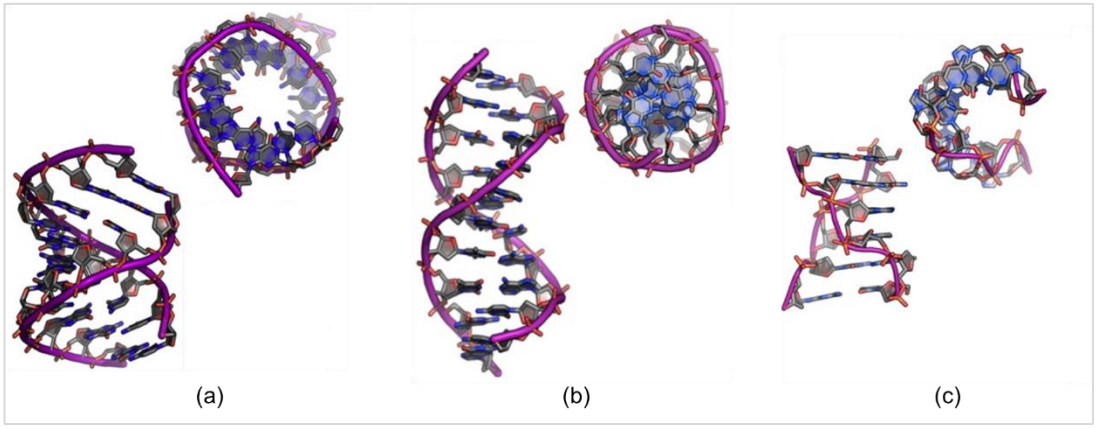


Figure 1.10: The three major types of DNA double helix, (a) A-DNA, (b) B-DNA and (c) Z-DNA.[39, 43-45]

1.4.1.2 B-DNA

James D. Watson and Francis Crick described the B-form DNA, which is believed to be in the majority of cells' nucleosome core.[46] In general, B-DNA is approximately 23.7 Å wide and can extend the helix by 34 Å per 10 base pairs of sequence. In solution, the double helix makes one whole turn around its axis for every 10.4-10.5 base pairs. It occurs in living cells at high levels of hydration.

1.4.1.3 Z-DNA

In 1979, a different type of double helical structure was discovered during the investigation of the DNA hexamer crystal structure of d(CGCGCG). The two strands were found to be joined by Watson-Crick base pairing in an antiparallel arrangement however, the helix was left handed.[44] Due to the unusual zigzag backbone of the structure, it was labelled as Z-DNA (**Figure 1.10**). The base pairs were flipped upside down because of the rotating cytidine moieties, introducing a criss-cross shape in the backbone leading to the conversion of B-DNA

to Z-DNA. Additionally, the guanine residues rotated to the *syn*-form from an *anti*-conformation (**Figure 1.11**); in contrast, the mononucleotide repeat observed in the B-DNA had a dinucleotide repeat as the nucleobases alternate in *syn*- and *anti*-conformations when moved along the Z-DNA sugar-phosphate chains. Alternate purine and pyrimidine moieties with d(CG) sequences readily form Z-DNA because the pyrimidine's preferred *anti*-conformation[47] requires less energy to form Z-DNA.

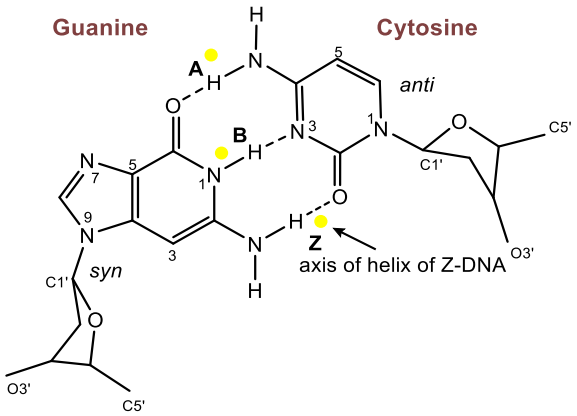


Figure 1.11: The helical axis of A-, B-, and Z-DNA.

Table 1.1: Geometric comparison of A-DNA[43], B-DNA[45] and Z-DNA[44]

Geometry attribute	A-form	B-form	Z-form
Helix sense	right-handed	right-handed	left-handed
Repeating unit	1 bp	1 bp	2 bp
Rotation/bp	32.7°	34.3°	60°/2
Mean bp/turn	11	10.5	12
Inclination of bp to axis	+19°	−1.2°	−9°
Rise/bp along axis	2.6 Å (0.26 nm)	3.4 Å (0.34 nm)	3.7 Å (0.37 nm)
Rise/turn of helix	28.6 Å (2.86 nm)	35.7 Å (3.57 nm)	45.6 Å (4.56 nm)
Mean propeller twist	+18°	+16°	0°
Glycosyl angle	Anti	Anti	pyrimidine: anti, purine: syn
Nucleotide phosphate to phosphate distance	5.9 Å	7.0 Å	dC: 5.7 Å, dG: 6.1 Å
Sugar pucker	C3'-endo	C2'-endo	dC: C2'-endo, dG: C3'-endo
Diameter	23 Å (2.3 nm)	20 Å (2.0 nm)	18 Å (1.8 nm)

1.4.2 Triplex DNA

Intermolecular triple helical DNA structures were first reported about 56 years ago[48-51] and were found to form at polypurine (R)-polypyrimidine (Y) tracts (**Figure 1.12**). A triplex DNA usually forms when a third strand binds to the major groove and binds to the purine rich strand of the underlying duplex DNA sequence by creating Hoogsteen and/or reversed Hoogsteen hydrogen bonds. There are two main scenarios for the triplex formation; when DNA with adjacent segments of homopurine and homopyrimidine runs can melt and refold to form intermolecular triplexes and binding of an oligonucleotide specifically to a DNA duplex.[52]

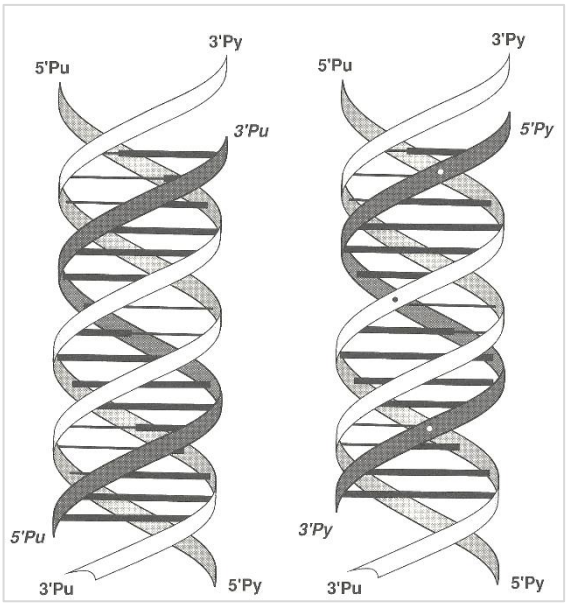


Figure 1.12: Diagram of DNA triplex structure.

Two basic forms of triplexes, R-R-Y and Y-R-Y, are formed depending on the type of nucleobases in the third strand which can be either purine (R) rich or pyrimidine (Y) rich. Reaction conditions also assist in the formation of these triplexes, as for example acidic pH allows the generation of Y-R-Y triplexes while neutral pH and divalent metal cations are the requirement of R-R-Y triplex formation in normal physiological conditions.[53] Triplex formation is mostly affected by the site of formation because of its structure, charge and rigid nature. Moreover, DNA sequence is another significant factor in triplex formation that needs to be either mirror repeats for intramolecular triplex, but for intermolecular triplexes the main requirement is that it should be a homopurine-homopyrimidine sequence.[54] The triple helical

DNA can be targeted selectively to control gene expression *in vivo* which presents the prospect of developing therapeutic agents like anti-cancer and anti-viral drugs. Triplex forming oligonucleotides (TFOs) also can be used to control gene expression, which can bind in the major groove of duplex DNA with high specificity and affinity. Because of these features of TFOs, it has been proposed as homing devices for genetic manipulation *in vivo*.[55]

1.4.3 Quadruplex DNA

Gellert and co-workers suggested quadruplex DNA by studying the X-ray diffraction configurations of gels produced by guanosine monophosphate (GMP) where four guanine residues formed a cyclic structure by forming hydrogen bonds in the centre.[37] This quadruple helical form of DNA is known as G-quadruplex (**Figure 1.13**), which forms from certain guanine-rich sequences. In 1958, four-stranded nucleic acid structures were first discovered during the study of the polyinosinic acid fiber. The X-ray diffraction pattern indicated the presence of parallel three- or four-fold polynucleotide strands,[56] which later on was proved to be quadruple-stranded.[48, 57]

The G-quadruplex structure is formed either by repetitive folding of the single polynucleotide molecule or by association of two or four guanine molecules. The structure consists of stacked G-quartets, which are square, co-planar arrangements of four guanine bases (**Figure 1.13**). G-quadruplex structures can be formed due to the addition of 1-, 2- or 4- guanine-rich DNA strands with different topologies (**Figure 1.13 c**), where the topology adopted by single-strands is possibly most significant for guanine-rich sequences at the 3′-ends of chromosomes(for example, telomeric G-quadruplex).[58]

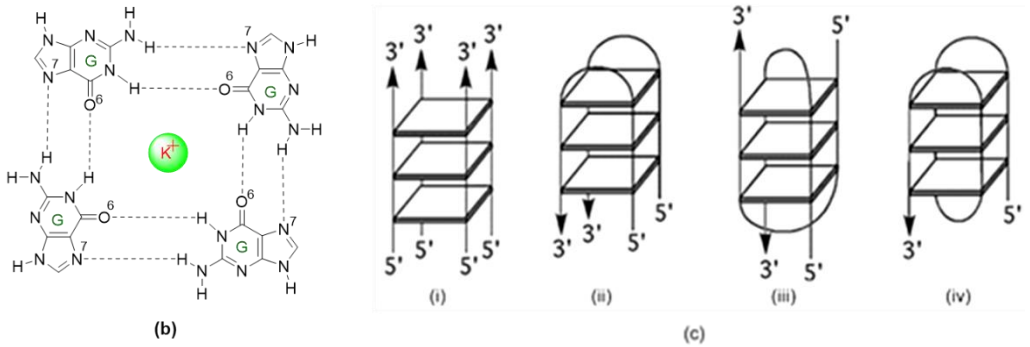


Figure 1.13: (**a**) G-quartet with central stabilizing cation, (**b**) Potential topologies for G-quadruplex structures, topologies constructed from four parallel strands (i), from two strands that are non-crossing (ii), cross-over (iii) and from a single strand (iv).[18]

The tetrad structure is stabilized through Hoogsteen hydrogen bond formation between the four guanine bases. These structures are reported to be found in telomeric regions along with the promoters of certain genes like c-Myc, c-Kit and Bcl-2 with individual structures. For example, telomeric DNA structures form parallel or antiparallel/parallel mixed-type structures in the presence of K^+ while antiparallel structures are formed in the presence of Na^+. On the other hand, c-Myc promoters form four-looped parallel structures, while in c-Kit promoters four connecting loops are found in three stacked G-quadruplex.[59] In 1989, Williamson and co-workers first reported the importance of monovalent cations (K^+, Na^+) in the formation and stabilization of the G-quadruplex structures and suggested that cations placed themselves in the cavity generated by the guanines of each quadruplex.[60]

A number of bioinformatic experiments have identified about 375,000 possible G-tetrad forming sequences in the human genome, though their existence has yet to be proved *in vivo*.[61] However, many indirect confirmations indicate the presence of G-quadruplex structures at several places in the human genome and they are suggested to be involved in various cellular activities. These quadruplex structures are found to control the expression of some genes that are involved in the progression of carcinogenesis and in the maintenance of telomere organization.[62] Their ability to stop telomeric elongation by inhibiting the telomerase

enzyme (which is found to be overexpressed in malignant cells) has prompted fantastic research interest.[18]

1.4.4 i-Motif

The i-motif is a tertiary DNA structure generated by intercalating two parallel duplex structures in a head-to-tail fashion where the strands are attached together through hemi-protonated C-C^+ pairs (**Figure 1.14**). This structure can be generated by the association of one, two or four strands containing four, two or one cytidine repeats, respectively. The centromeric and telomeric C-rich repeats can fold to form an intramolecular i-motif structure.[63, 64]

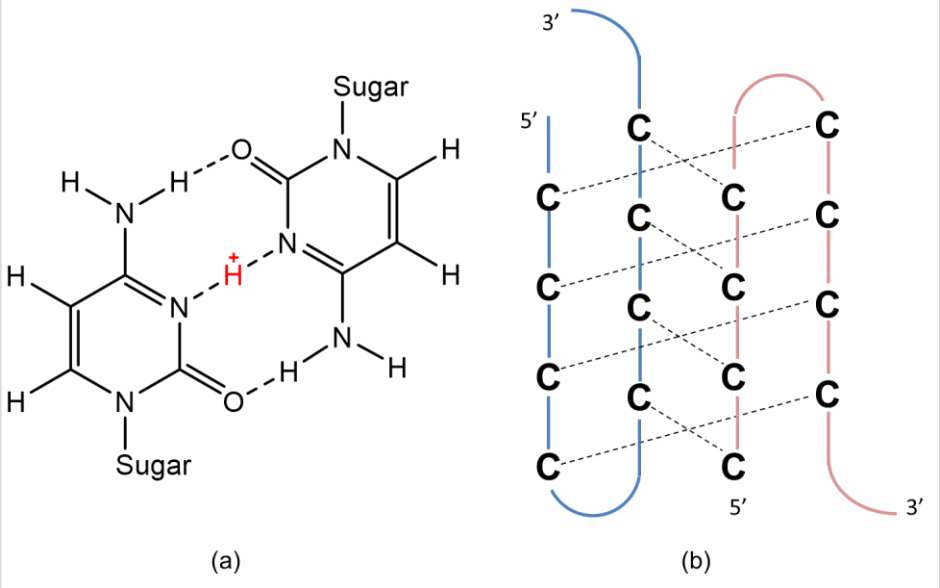


Figure 1.14: (a) Hemi-protonated C-C^+ pairs (*note, hydrogen bonds are not of an appropriate length*) (b) Intermolecular i-motif pattern.

Slightly acidic pH favours and guides the interconversion of the C-rich ss-DNA into an i-motif and forms an antiparallel structure.[63] Moreover, the pH-dependent structure is found to be significantly affected by the cytosine base numbers, environmental conditions and the length of the loops.[65] It is well known that >40% of all genes contain C-rich sequences in or near the regulatory regions, including the oncogenes.[66]

1.5 DNA Major and Minor Grooves

The DNA double helix is a reasonably rigid and viscid molecule of huge length and small diameter, which has spaces, or grooves (**Figure 1.15**), between the strands. Together with the base pairs, these empty spaces are possible binding pockets or sites for proteins or drugs. Such grooves are unequally sized and classified into two types, the major groove which is 22 Å in width and the minor groove which is 12 Å wide. The major groove is deep and wide, whereas the minor groove is narrow and shallow.[67] The major and minor grooves are reverse to each other, and each runs constantly along with the full length of the DNA molecule; usually such grooves form due to the antiparallel double strand arrangement. It has been established that these grooves play a vital role by providing space for a protein to bind for cell cycle processes like transcription and DNA replication. Hydrogen bonding, van der Waal's interactions and electrostatic interactions play key roles in protein-DNA binding. Thus it is hypothesised and proven in many cases that small molecules can also bind to the major as well as minor grooves, resulting in a distorted DNA structure or interruption of protein-DNA binding, followed by a targeted beneficial effect (for example, anticancer agents, HIV treatment).

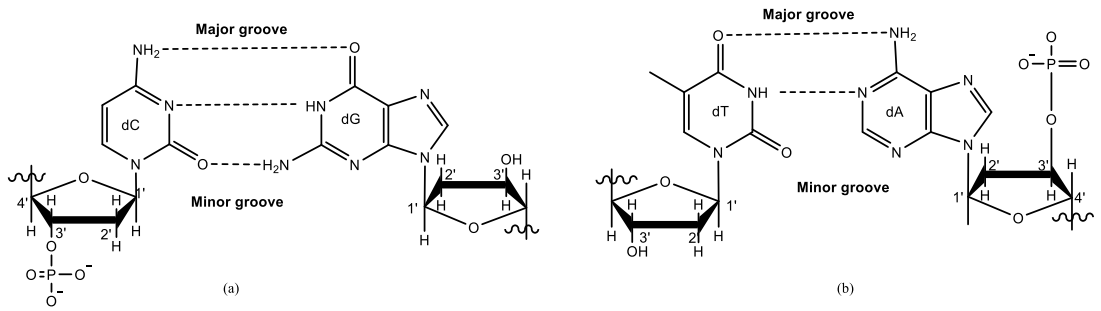


Figure 1.15: (a) Major and minor groove formation at a dC-dG base pair and (b) major and minor groove formation at a dT-dA base pair.

The widths of the major and minor grooves are defined by the distance of the phosphate-to-phosphate backbone through two strands in a direction perpendicular to the helical axis of the strands. In general conformation, a wide major groove is observed in B-DNA, while a narrow major groove is present in A-DNA that leads to a wider minor groove in A-DNA.[39, 68]

1.6 Telomeric DNA

The idea of telomeres originated in the late 1930s. Muller and McClintock autonomously observed that regular chromosome ends have special properties to protect them from end-to-end fusion using cytogenetic tactics. After extensive study in the past few years, it has been confirmed that highly preserved telomere structures at the ends of linear chromosomes with tandem repeats of DNA sequences and associated proteins are the key components for conserving DNA structure.

The telomere is responsible for some vital biological functions (**Figure 1.16**), for example, protecting chromosomes from recombination, end-to-end fusion and recognition as damaged DNA; providing resources for complete replication of chromosomes; contributing to chromosomal organization within the nucleus; participating in gene expression regulation; and also serves as a molecular clock by controlling the replicative ability of human cells and their entry into senescence.

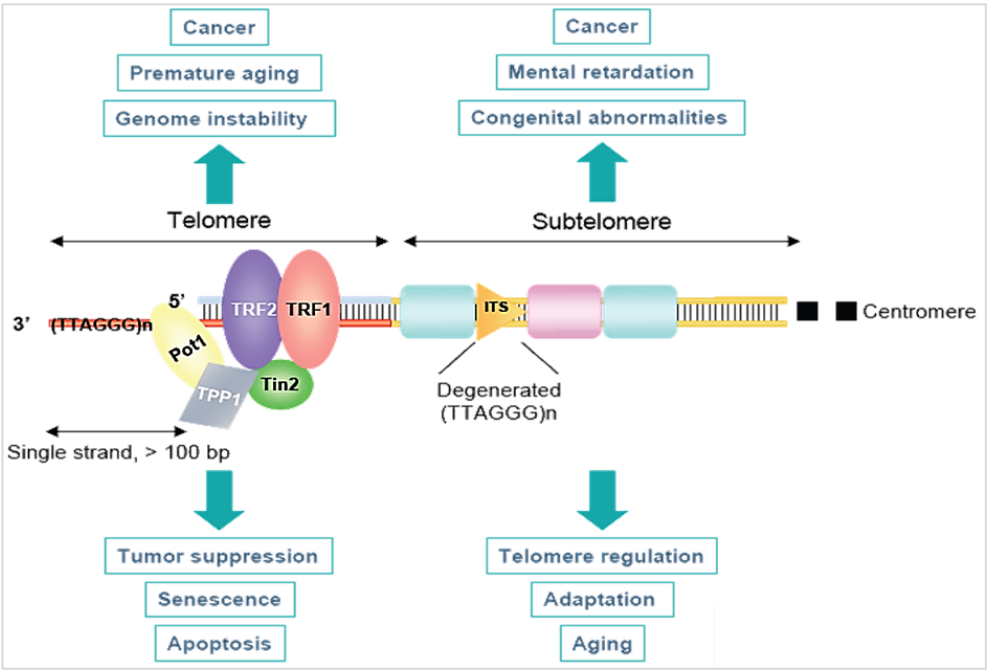


Figure 1.16: The consequences that arise from telomere function and dysfunction.[69]

Replication of chromosome ends presents a distinct problem for cells. Where DNA replication by polymerase proceeds in a 5′ → 3′ direction, Watson and Olovnikov individually postulated that DNA is gradually lost from the ends of chromosomes during each instance of cell division, as the conventional DNA polymerase cannot fully replicate the 3′-end of the lagging strand. This is known as the end replication problem. The newly synthesized leading strand cannot generate overhangs; however, the lagging strand loses its extreme 3′-end after RNA primers are removed. (**Figure 1.17**).

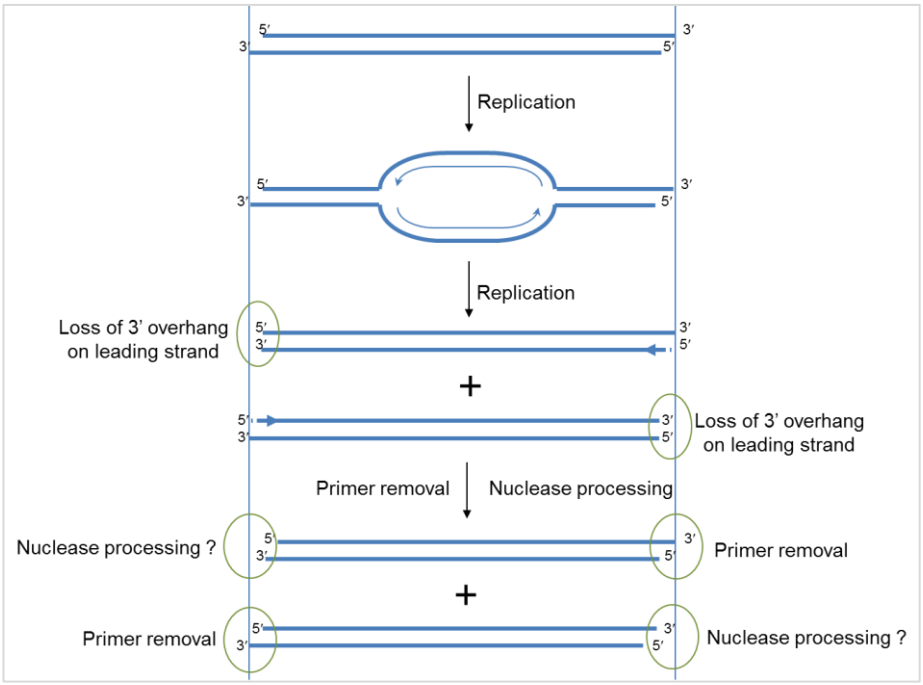


Figure 1.17: Schematic representation of the end replication problem.

Such telomere shortening is detected with progressive cell division *in vitro* and enlarged *in vivo*.[70-72] In general, regular mammalian somatic cells multiply a limited number of times *in vitro*, where the maximum number of cell proliferation cycles is referred to as the Hayflick limit[73], which acts as a molecular clock. At the Hayflick limit, one or more critically shortened telomeres initiate a permanent growth arrest known as replicative senescence or mortality stage 1 (M1).[74, 75] If cells can escape M1 due to inactivation of critical cell cycle checkpoint genes (such as p53) then those cells continue to divide and suffer more telomere loss until reaching the second proliferative block, crisis, or mortality stage (M2)[76-78], which is categorised

by massive cell death caused by critically short and dysfunctional telomeres. Rare survivor cells escaping from crisis are able to maintain telomere length, in most cases by activation of telomerase, and this leads to unlimited proliferative capacity, *i.e.*, cellular immortalization.

Telomerase, a reverse transcriptase enzyme, is a RNA-dependent DNA polymerase that synthesizes telomeric DNA repeats at 3′-overhangs. In 1985, the enzyme was first discovered in a protozoa, *Tetrahymena thermophila*[79], and interestingly telomerase activity was also found in over 90% of malignant cells and *in vitro* immortalized cells. However, it is inactive in most normal human somatic cells.[80, 81] In 1988, Moyzis and co-workers reported that telomeres consist of a tandem series of guanosine-rich nucleic acid hexamers [$(TTAGGG)_n$] and various associated binding proteins, forming a structure that preserves chromosomal stability.[82] The termini of telomeres consist of 25-200 nucleotides of single-stranded DNA, usually a G-rich strand (referred to as the "G-overhang"). It is presumed that most of the remainder of the telomere forms duplex DNA by Watson-Crick base pairing, but the G-rich strand is capable of Hoogsteen base pairing and forms planar G-quartet structures that stack on each other to form G-quadruplexes.[83]

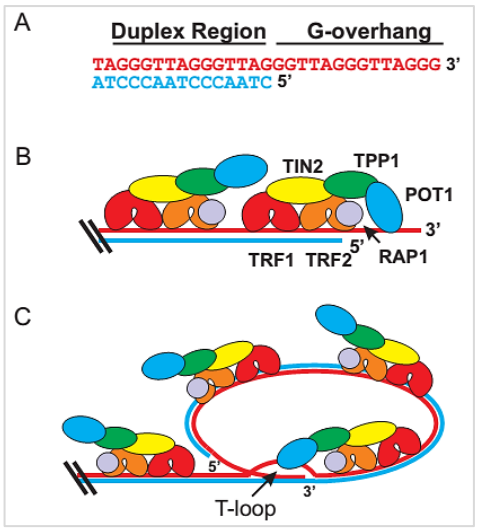


Figure 1.18: Telomere structure: (A) Telomeric DNA consists of a repetitive DNA sequence of a duplex region and single-stranded DNA G-strand overhang. (B) The shelterin complex binds to both the duplex and single-stranded DNA regions through specific protein/DNA interactions. (C) Formation of t-loop involving strand invasion of the G-overhang to create a displacement-loop (D-loop).

The ends of telomeres form a lariat-like structure called a t-loop (**Figure 1.18**)[84], which is assumed to be formed by strand invasion of the 3′ single strand overhang into the earlier double stranded telomeric DNA that is stabilized by an attached shelterin protein (telomeric protein complex).[85] This protein complex maintains telomere length, promotes t-loop formation, recruit telomerase to telomeric ends and thereby protects the ends of chromosomes from DNA damage.[86-88]

TRF1 and TRF2 are the duplex telomeric binding proteins which are specifically localized to the telomeric DNA sequence. TRF1 and TRF2 and their accompanying proteins are primarily responsible for stabilizing the complex and forming the t-loop. Such stabilization is essential to form the quadruplex structure due to the G-rich nature of the telomere overhang. This telomeric repeat binding factor, TRF1, is important for intratelomeric winding.[84] Similarly, TRF2 binds beside the telomere sequence and is important for stabilizing and forming the t-loop as TRF2 is mainly abundant at the base of the t-loop.[84] Thus TRF1 and TRF2 work together as TRF1 forms a loop and TRF2 tightens up the strand and locks it.

Additionally, duplex telomeric DNA binding proteins have some self-associated proteins.[89] For example, Human Rap1 (hRap1) protein is combined into the t-loop and interacts with TRF2, however there is no reported specific role of hRap1 in humans.[90] Shay also reported that tankyrase, TRF1-interacting ankyrin-related ADP-ribose polymerase enzyme, releases from the t-complex and can inhibit TRF1. On the other hand, TIN2 stimulates TRF1 function and causes it to bind to the telomere.[89] The DNA damage response complex, RAD50/MRE11/NBS1, also functions along with TRF2, where MRE11 usually functions in homologous recombination to repair the disruptions in the DNA double strand.[91] It is believed that the single stranded tail enters the duplex telomere to stabilize the d-loop at the telomere and based on this function *in vitro*, NBS1 plays the role of unwinding the t-loop *via* a helicase during the S phase of the cell cycle.[92]

1.6.1 Functional Assembly of Human Telomerase

Telomerase is a reverse transcriptase enzyme consisting of a RNA template complementary to the telomeric hexameric repeats (TTAGGG) at the chromosome end, which is responsible for lengthening and maintaining the length of the telomeres and subsequent extension of the number of cell divisions.[93] The telomerase enzyme is composed of a RNA subunit (hTR), a protein subunit (hTERT) and other accompanying proteins. In 1992, Blackburn first reported that the reverse transcriptase enzyme catalyses the addition of DNA bases (TTAGGG) to the telomeric end (3′) that corresponds to the RNA template sequence of hTR, and later in 2001 Wright and co-workers also confirmed the similar functions of hTR.[94, 95] Two other proteins, p23 and hsp90, are also necessary *in vivo* to assemble the telomerase components in the telomerase complex and this was confirmed *in vitro* by using recombinant hTR and hTERT that were sufficient to reconstruct the enzyme complex.[96]

1.6.1.1 RNA Subunit, hTR

It is already known that repetitive telomere sequences cover the ends of eukaryotic chromosomes and protect them from destruction and rearrangements. These telomeric repeats are synthesized by the telomerase RNA-protein complex, where the duplicating of the RNA component of human telomerase is termed hTR. The RNA part (hTR) of telomerase provides the template for the telomeric repeat synthesis. In humans, RNA polymerase-II transcribes hTR and hTR is processed at the telomeric 3′-end to produce a mature transcript of 451 nucleotides.[97] The template for reverse transcription lies close to the 5′-end of the sequence (between nucleotides 46 and 53), and specifies telomeric DNA sequence synthesis. The hTR template region is comprised of 11 nucleotides (5'-CUAACCCUAAC) corresponding to the hTelo sequence $(TTAGGG)_n$. In general, malignant cell lines and germline tissues express higher amounts of hTR compared with normal somatic cells and tissues.

1.6.1.2 Protein Subunit, hTERT

hTERT, or Human Telomerase Reverse Transcriptase, is another vital part of the telomerase enzyme. In general, telomerase does not either express or activate during each instance of mitotic cell division that results in telomere shortening. However, telomerase can extend telomeres in critical condition, which is essential for cells to escape from senescence. hTERT is highly expressed in malignant tumours and cancer cell lines but not in normal tissues or telomerase-negative cell lines.

It has been reported that several transcription factors and oncogenes (*e.g.*, c-Myc, Sp1, HIF-1, AP2) can activate hTERT, whereas many cancer suppressing genes (*e.g.*, p53, WT1 and Menin) can initiate several factors with the ability to suppress hTERT activity.[98, 99]

1.6.2 Telomeric DNA/RNA Hybrid Duplex

According to Shaw and co-workers, the DNA/RNA hybrid duplex was first proposed six years after the double-helical structure of DNA was reported by Watson and Crick.[100] In 1960, the first DNA/RNA hybrid duplex structure was synthesized by reacting oligodeoxythymidylic acid with polyriboadenylic acid.[101] The first DNA/RNA hybrid duplex was formed through the annealing of an RNA strand with a complementary DNA strand in 1961.[102]

Such hybrid duplexes composed of a DNA and a RNA strand are initiated in various important biological processes, for example, hybrids are formed as an intermediate during transcription, DNA replication, and reverse transcription process by retroviruses.[103]

In 1967, x-ray diffraction studies and CD spectroscopy confirmed the difference of the hybrid conformation from than of B-form DNA.[104] Though the crystal structure of a DNA/RNA hybrid duplex was solved in 1982, hybrid formation during the replication process was discovered in 1975, making it an appealing drug target.[100]

Telomeres, the ends of every chromosome, exist as a protein-attached DNA composite with long repeats of a unique six base sequence (5′-TTAGGG). Upon activation of telomerase, hTERT can add the same six-nucleotide repeating sequence, 5′-TTAGGG, to the 3′ strand of the chromosome by using TERC (where the template region of TERC is 3′-CAAUCCCAAUC-

5′).[105] The synthesis of a telomeric repeat involves the 3′-end of the DNA primer base pairing with the RNA template to form a DNA/RNA hybrid duplex positioned within the active site. DNA polymerization then proceeds *via* reverse transcription of the RNA template to synthesize one telomeric repeat (TTAGGG) on the 3′-end of the DNA primer. The steps of telomerase activity and DNA/RNA hybrid duplex formation are as follows (see **Figure 1.19**)[106]:

A. A schematic diagram of the human telomerase enzyme bound to the telomeric DNA substrate. Prior to repeat synthesis, the hTR alignment region (green) base pairs with the 3′-end of the telomeric primer DNA (blue) to form a 5-bp DNA/RNA hybrid duplex within the active site. The RNA template is controlled by flanking sequences bound to the TERT protein, while the 5′-region of the hTelo DNA is bound to the TEN domain. The TERT protein catalyses polymerization of six deoxyribonucleotides to the 3′-end of the DNA primer by reverse transcribing the hTR template sequence (orange) (step 1).

B. After nucleotide addition, a new repeat (red) is produced, while the DNA/RNA hybrid duplex length remains constant at the initial 5-bp. Upon getting to the end of the template, the DNA/RNA duplex dissociates from the active site (step 2).

C. Then the DNA/RNA hybrid duplex undergoes strand dissociation (step 3a), and the DNA reanneals to either the alignment or the template regions. If the DNA reanneals to the template region, the result is non-productive, requiring further strand separation and a new attempt at reannealing. If the DNA reanneals to the alignment region (step 3b), the RNA template is regenerated. Both strand separation and realignment are presumably reversible, as indicated by the double-headed arrows.

D. The realigned DNA/RNA hybrid duplex is then recaptured by binding to the active site (step 4) for further nucleotide addition. Unsuccessful reannealing of the DNA primer to the RNA template results in complete dissociation of the product from the enzyme (step 5).

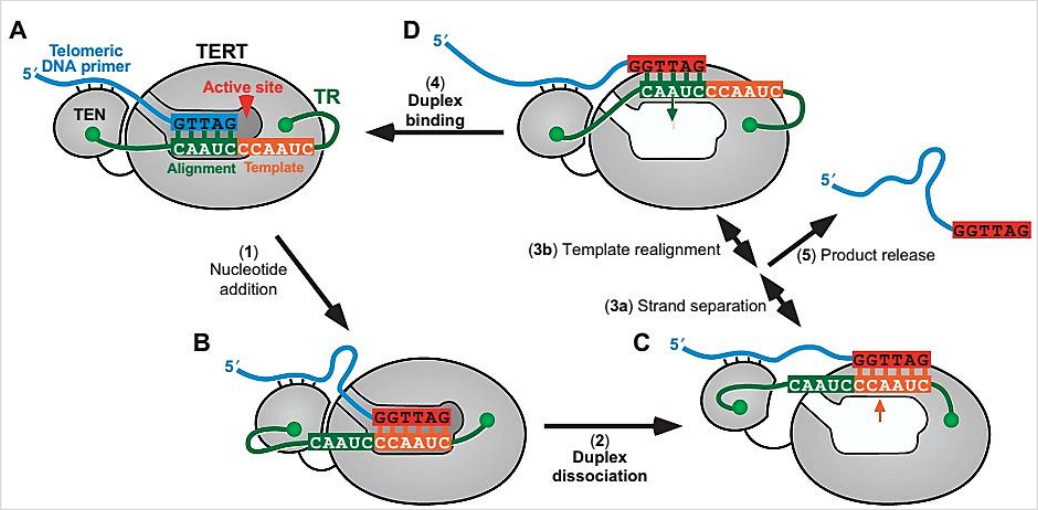


Figure 1.19: Schematic of a functioning model for telomerase activity.[106]

Hence, binding of small ligands to the DNA/RNA hybrid duplex can theoretically offer a therapeutic method to control telomerase activity by preventing telomere extension and/or by altering the substrate/enzyme interaction, and/or stopping dissociation of the enzyme from the substrate.[107]

DNA/RNA hybrid structures are also formed during different genomic functions rather than telomerase activation. DNA mediated mRNA synthesis is a set process for protein synthesis which initiates at promoter sites within the DNA template, where RNA polymerase generally binds to the initiation site, unwinds the helical turn of duplex DNA and starts sequential addition of ribonucleotides complementary to the DNA bases. Termination of this process accomplishes with the release of the RNA strand from an unstable DNA/RNA hybrid.[108] During the reverse transcription process, the RNA mediated synthesis of DNA is an integral part of the retroviral life cycle, which was observed in 10% of the eukaryotic genome as a direct product of reverse transcription[109]. Reverse transcription is similar to the normal DNA transcription process, but the reverse transcriptase enzyme is engaged here to reverse transcribe the DNA from a RNA template.[109, 110] DNA/RNA hybrid formation can be observed during reverse transcriptase mediated DNA synthesis and prior to reverse transcriptase mediated RNaseH cleavage of the template RNA strand. DNA/RNA hybrid duplexes play an important role in the DNA replication process, which begins with the unwinding of duplex DNA

by DNA helicase followed by priming of the DNA strands by proteins and enzymes[111] and proceeds in leading strand and lagging strand synthesis. Lagging strand synthesis utilizes Okazaki fragments (*e.g.*, in eukaryotic cells)[112], where DNA/RNA hybrids of 100-200 nucleotides are formed in the presence of a RNA primer, which is later degraded with the help of DNA polymerase and re-sealed using DNA ligase. DNA/RNA hybrid duplex formation is essential to initiate the transcription process within mitochondrial DNA and such hybrids serve as primers for DNA replication of the leading strand.[111, 113]

1.6.2.1 Structure and stability of DNA/RNA hybrid

An extensive study and literature review of DNA/RNA hybrid stability suggested that by varying the percentage of $(A)_n$:$(T \text{ or } U)_n$ content, oligomeric length and the percentage of deoxyribo-pyrimidine (d(Py)) bases in each of the strands, the stability of such a hybrid could be estimated. Lesnik and Freier demonstrated that, typically, an inverse relationship is observed between the percentages of (A)n:(T or U)n content in the hybrid duplex and the thermal stability; hybrid duplexes of equal d(Py) and d(Pu) content are thermally similar when comparing the d(Pu):r(Py) and d(Py):r(Pu) forms; decreasing oligomeric length decreases the thermal stability; isolation of (A)n: (T or U)n content to continuous tracts results in a lower thermally stability; and continuous tract (A)n:(T or U)n content results in d(Py):r(Pu) thermal stability lower than that of the corresponding d(Pu):r(Py) form under physiological conditions.[114]

DNA/RNA hybrid stability is also related to the ability of the hybrid duplex to associate with a third single strand and it has been observed that the d(Py):r(Pu) hybrid cannot accommodate a third strand of DNA, while d(Pu):r(Py) hybrids can be associated with a third strand.[115] Moreover, the stability of hybrid duplexes is linked to the relative stability of the ribose chain because the unfavourable van der Waal's forces experienced between the sugar and pyrimidine bases can result the purine nucleotide being highly stable compared with the pyrimidine nucleotide.

Different researchers resolved the DNA/RNA hybrid structures by using a variety of hybrid duplexes consisting of polypyrimidine and polypurine tracts of both DNA and RNA bases. In a brief, studies of the crystal structures of chimeric DNA/RNA hybrids designated them as pure

A-form in conformations like RNA duplexes.[116, 117] However, NMR analysis of the DNA/RNA hybrid structures in solution showed that the DNA strands of hybrid duplexes have helical conformations which are transitional between the A- and B-forms DNA.[118-120] Researchers also tried to define the stability of various forms of DNA/RNA hybrid duplexes and it has been reported that the r(purine)-d(pyrimidine) hybrids are more stable than the hybrids formed with r(pyrimidine)-d(purine).[121-124] The polymeric DNA/RNA hybrid structure was also studied to identify the conformational alterations upon changes in relative humidity. Milman and co-workers showed that in extreme solvated conditions poly(rA):poly(dT) structure is similar to that of B-form DNA duplexes.[125] However, further analysis by solid state NMR recommends that the same hybrid duplex (*i.e.*, poly(rA):poly(dT)) is able to exist as an A-like conformation at relatively low humidity.[126] Equilibrium Raman spectroscopy later explained the conformational differences occur due to alterations in furanose sugar conformation in RNA and DNA strands, where the over-all B-like conformation is dictated by the thymidine methyl in the d(Pyrimidine) strand.[115]

The ribose sugar pucker is the major participant in determination of the helical parameters of nucleic acids. In general, the ribose sugar pucker of RNA is found in the C3′-endo conformation, which facilitates the conformation of homoribonucleic strands to A-form conformation. On the other hand, the deoxyribose conformation is much more organised to a number of different conformations and usually consists of a C2′-endo (or equivalent C3′-exo) conformation, which is characteristic of the B-form. Another sugar pucker is observed as an intermediate between the C3′-endo and C2′-endo conformations, which is termed as C4′-endo and arises at various points in the subsequent formation of DNA/RNA hybrids helical structures.[127]

In 1999, Conn and co-workers resolved the crystal structure of the DNA/RNA hybrid composed of d(GCTTCTCTTC)/r(GAAGAGAAGC). They have calculated the average helical parameters of the double helix (using the program NEWHEL93, R.E.Dickerson, UCLA), which were 30.4° for helix rotation and 2.9 Å for helical rise. Thus 11.8 nucleotides were accommodated per complete helical turn and the pitch of the helix was 34.2 Å. These values confirm that the hybrid duplex belongs to the A-form DNA structure, which is also consistent with the A-form

RNA duplex structure. All deoxyribose moiety, except C20, does not adopt a sugar pucker adjacent to C3′-*endo,* which is believed to be inclined by crystal packing contacts, which is typical of the the A-form structures. Moreover, the geometry of the minor groove of the hybrid duplex also bore a resemblance to the characteristics of the crystalline A-form double helices, such as the average width being 9.8 Å, which is very close to the average width of typical crystal structures of hybrid duplexes found in the Nucleic Acid Database.[128]

1.6.3 Telomeric G-quadruplex Structure

In several studies, it was described that telomeric DNA is capable of forming a G-quadruplex structure in biological systems. Single-stranded telomeric terminal sequences from a number of organisms were found to form G-G paired hairpin type structures,[129] which later on were defined in *Tetrahymena* as telomeric DNA.[130] *Stylonychia lemnae* telomeric DNA was used to establish the biological existence of the quadruplexes.[131] Recently it has been evidenced that telomere end-binding proteins, TEBPα and TEBPβ, control *in vivo* G-quadruplex formation because it was confirmed later that phosphorylation of TEBPβ during replication is essential to resolve the quadruplex structures.[132] *In vivo* G-quartet formation was further proved by the binding and detection of the radiolabelled ligand 360A (2,6-*N,N′*-methylquinolinio- 3-yl-pyridine dicarboxamide) and/or the fluorescent ligand BMVC (3,6-*bis*(1-methyl-4-vinylpyridinium) carbazole diiodide. Researchers not only proved the existence of hTelo quadruplex structures, but they are also found evidence of the existence of such quadruplexes in the non-coding regions of genes, particularly in gene promoter sites. In 2010, it was reported that such quadruplexes are present near the transcription start sites of promoter regions which strongly supports the link between nuclease-hypersensitive elements (NHE) and quadruplex-forming sequences.[133]

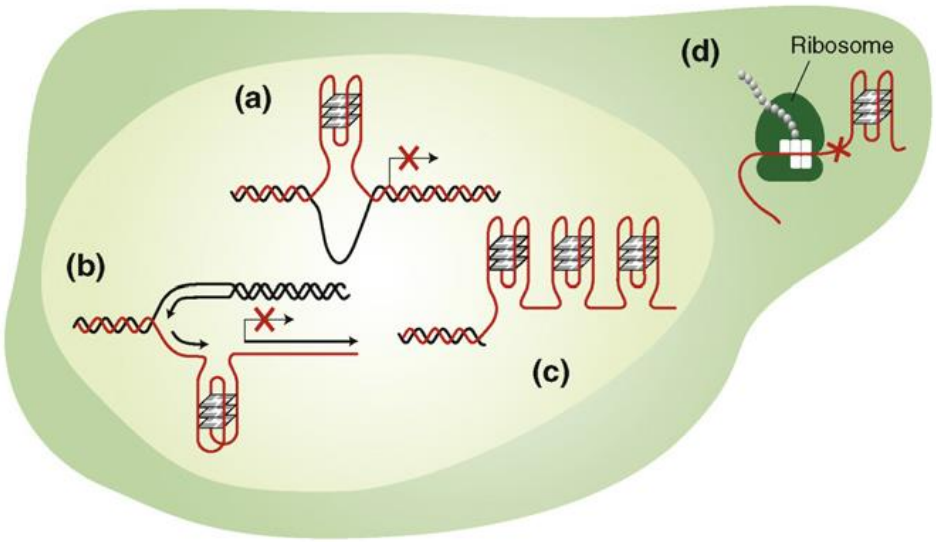


Figure 1.20: Probable places of the G-quadruplex structures within cells: in nucleus (a) during replication (b) transient single stranded form of DNA (c) telomeric single stranded G-rich overhangs; outside nucleus (d) G-quadruplexes in mRNA. (G-quadruplex mediated blockade during replication, transcription and translation labelled by red cross).[134]

In an earlier study with *Oxytricha nova,* it was observed that folding in telomeric DNA and formation of quadruplex structures can block the elongation of telomeres by the telomerase enzyme and the result of this study demonstrated that interconversion of the telomeric DNA into the G-quartet structure causes delayed elongation of the guanine strand by telomerase (**Figure 1.20 a**).[135] Moreover, several research laboratories have reported stabilization of hTelo G-quadruplex structures through using various groups of drugs that can inhibit telomeric elongation by telomerase (**Figure 1.20 b**). Reported studies demonstrated that drugs that can bind and stabilize the G-quadruplex structures of telomeric sequences can also block *in vitro* action of telomerase as well as serving to inhibit telomere elongation *in vivo* and the proliferation of cancer cells (**Figure 1.21**).[136-138]

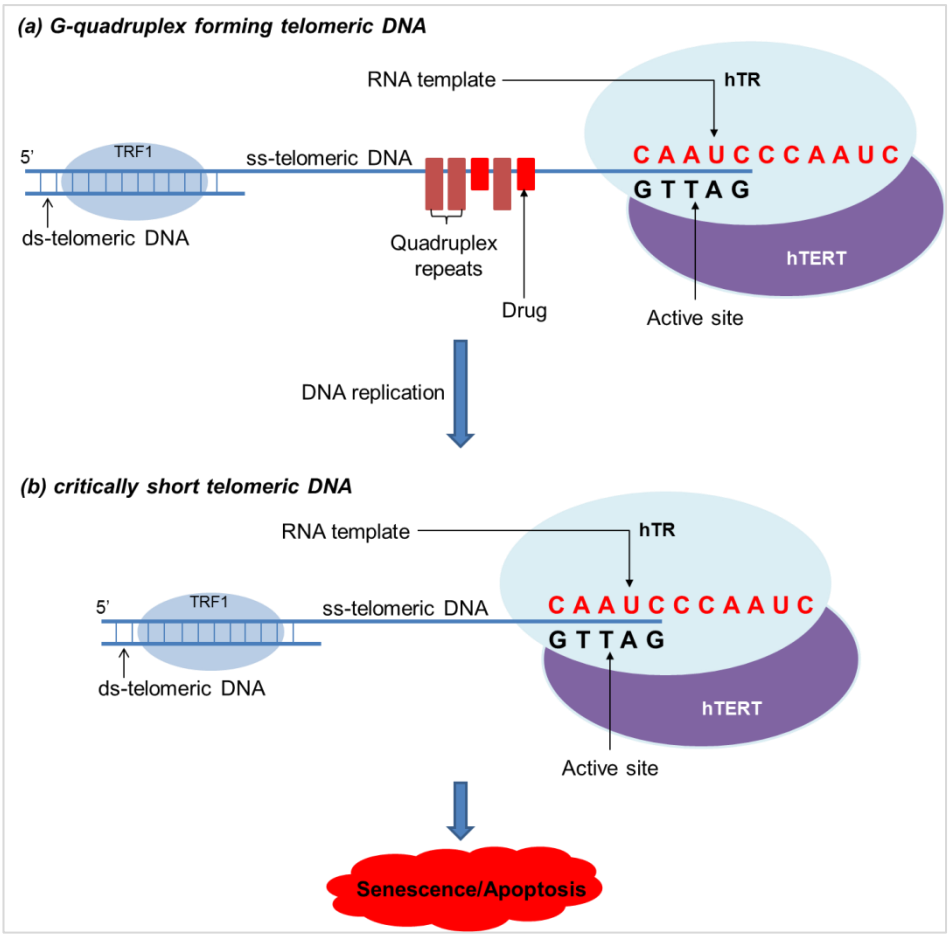


Figure 1.21: Telomere length regulation (a) Inhibition of the telomerase enzyme by G-quadruplex formation at the 3′ end, which was further stabilized by ligand interaction inhibiting further telomere elongation. (b) Diagram showing the effects of inhibition of telomere elongation; critically formed hTelo that ultimately leads to senescence/apoptosis.[139]

In a recent study in 2014, Komiyama and co-workers showed that telomeric DNA is transcribed into telomeric repeat-containing RNA (UUAGGG repeats) in mammalian cells.[140] It has also been reported that the two-repeats of human telomeric sequence, $d(TAGGGT)_2$, can form a dimeric parallel-stranded G-quadruplex structure, whereas a three-repeat, $d(TAGGGT)_3$, and a single-repeat, d(TAGGGT), of the human telomeric sequence form a (3+1) dimeric DNA G-quadruplex structure.[141,142] Marin and co-workers showed the formation of a RNA/PNA hybrid G-quadruplex structure with RNA and PNA.[143] Moreover, it is suspected that telomere RNA can bind to telomere DNA by forming an intermolecular DNA/RNA hybrid G-quadruplex (**Figure 1.22 c**), because of difficulties in its identification by typical spectroscopic tools like

NMR and crystallography. This is because DNA G-quadruplexes, RNA G-quadruplexes and DNA/RNA hybrid quadruplexes can exist together as a mixture in solution.

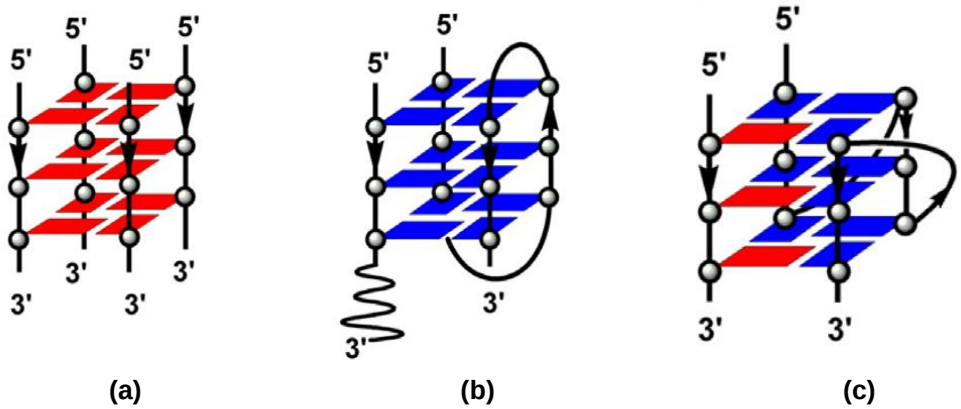


Figure 1.22: Illustration of three possible G-quadruplex structures that may exist together as a mixture in solution, where (**a**) tetramer RNA-quadruplex can form by 6-mer hTelo RNA sequence, (**b**) dimeric DNA-quadruplex can form by 16-mer hTelo DNA sequence and (**c**) DNA/RNA hybrid-quadruplex can form by 6-mer RNA and 16-mer hTelo DNA.

NMR studies have already resolved the usual sequence of hTelo G-quadruplex structures (**Figure 1.23**) that are formed by 22-mer hTelo DNA (effective four TTAGGG repeats) in Na^+ solution and exhibit an *anti*-parallel topology with one diagonal and two lateral loops.[144] However, the crystal structure of the same sequence in K^+ solution showed a complementary all-parallel topology with three strand-reversal loops. The sequence $d(TTAGGG)_2T$ was also solved by NMR spectroscopy and X-ray crystallography in K^+ solution, where both experiments revealed a bi-molecular parallel topology.[141]

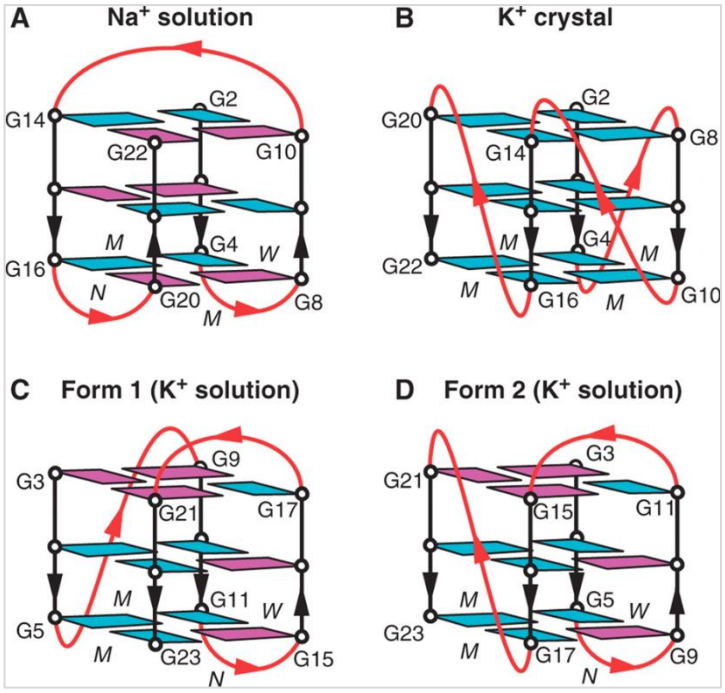


Figure 1.23: Different forms of intramolecular quadruplex formation by the hTelo repeat sequence (d[AGGG(TTAGGG)$_3$]), where loops are coloured **red**, and *anti*- and *syn*-guanines are coloured **cyan** and **magenta**, respectively and **W**, **M** and **N** denote **wide**, **medium** and **narrow** grooves, respectively.[145]

Other variations in the telomeric sequence have shown different structures as determined by NMR. Some of these have resulted from the modification of the sequence in order to solve the structure. The wide variety of topologies (**Figure 1.24, 1.25**) published for sequences related to the hTelo DNA sequence indicate the dynamic and heterogeneous nature of the human telomeric G-quadruplex.

More recently, many of the structures published have shown a (3+1) topology (**Figure 1.26 c**). The (3+1) quadruplex topology covers those distinctive patterns of (3+1) *syn*- and *anti*-glycosidic angles at each G-quartet.

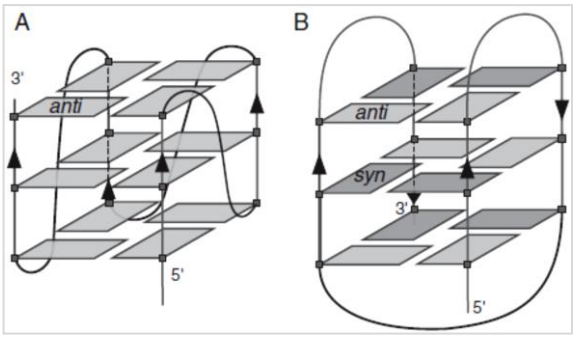


Figure 1.24: Folding patterns of the hTelo sequence (22-mer); (A) propeller-type parallel-stranded intramolecular G-quadruplex formed in K^+ solution, (B) basket-type mixed parallel/antiparallel-stranded intramolecular G-quadruplex in Na^+ solution.

The (3+1)-formed quadruplex topology was first observed in an unequal structure containing a three-repeat sequence associated with a separate single-repeat sequence. For instance, the sequence d[TTGGG(TTAGGG)$_3$A] exists as a 95% single species in K^+ solution, which was confirmed by NMR studies of the (3+1) intramolecular arrangement with a single chain-reversal loop followed by two lateral loops at each end of the quadruplex structure. This is also observed in the solution structure of d[A$_3$G$_3$(TTAGGG)$_3$AA], indicating the significant stabilizing role of neighbouring sequences.

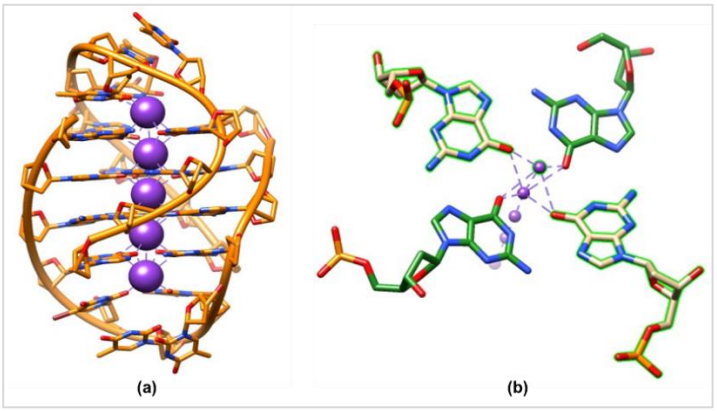


Figure 1.25: The crystal structure of the bimolecular quadruplex formed by the *Oxythricha nova* telomeric sequence d(G$_4$T$_4$G$_4$) (PDB entry 1JPQ); (a) overall topology is indicated by the orange ribbon with purple spheres of K^+ ions, and (b) a projection down the central passage representing the relative widths of the four grooves.[146]

A similar result was also observed for the NMR structure of d($AG_3(TTAGGG)_3$) in K^+ solution. Most commonly, the structure is of three G-quartets connected with a double-chain-reversal side loop and two lateral loops, each consisting of 3 nucleotides, TTA.[147]

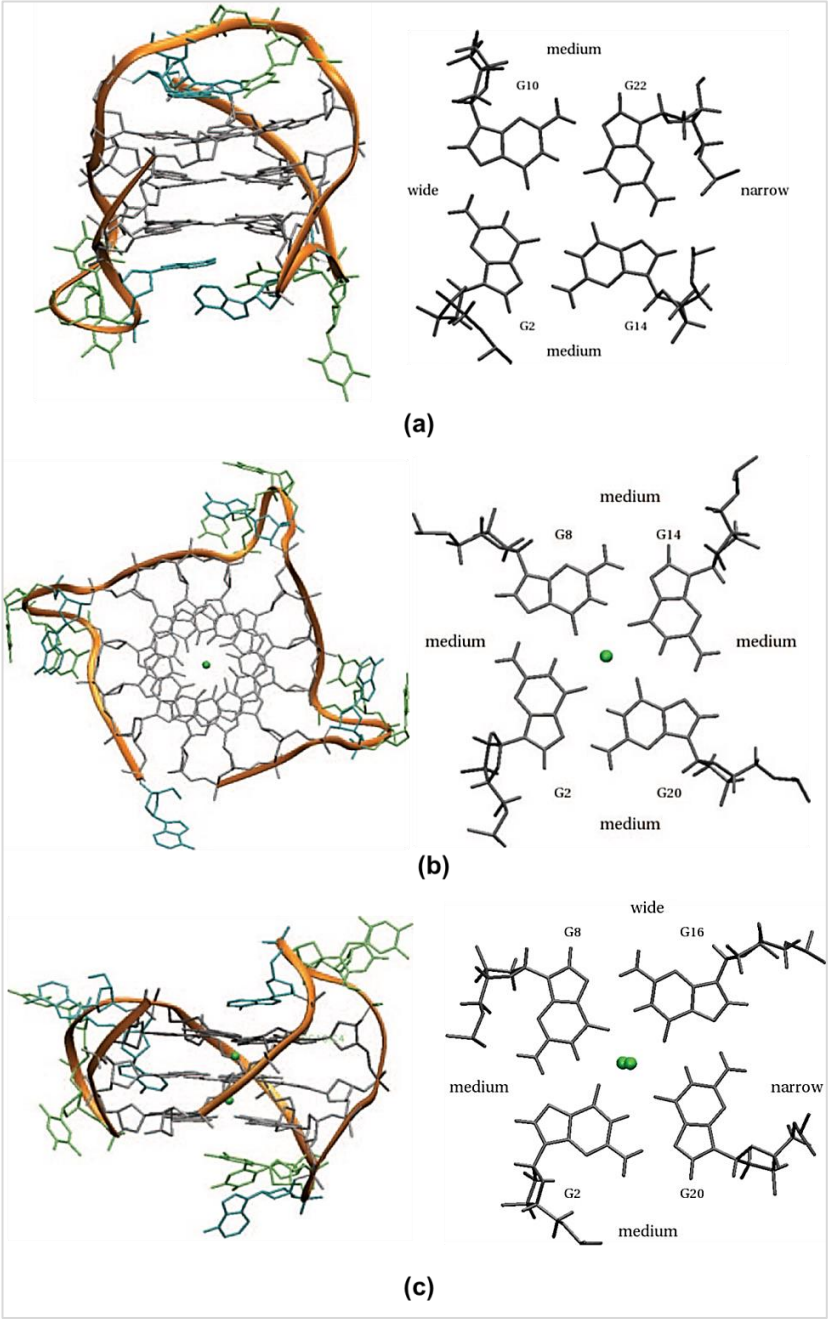


Figure 1.26: Structures of the hTelo quadruplex for the d[$AGGG(TTAGGG)_3$] sequence, (a) in Na^+ solution with one diagonal and two lateral loops (PDB entry 143D)[144], (b) in K^+ solution that forms A-DNA with three strand-reversal loops (PDB entry 1KF1)[141], (c) in K^+ solution, that forms B-DNA with one strand-reversal and two lateral loops, as determined by NMR.[147,148]

A second form of (3+1) structure results from the two lateral loops followed by one side chain reversal loop. This form has mainly been seen in structures where the 3′ end of the flanking sequence has been altered to include TT. One such structure is that of the NMR structure of the sequence d(TAGGG(TTAGGG)$_3$TT) in K^+ ion solution.[148] In K^+ ion solution these two forms have been found to exist in equilibrium (**Figure 1.27**).[133]

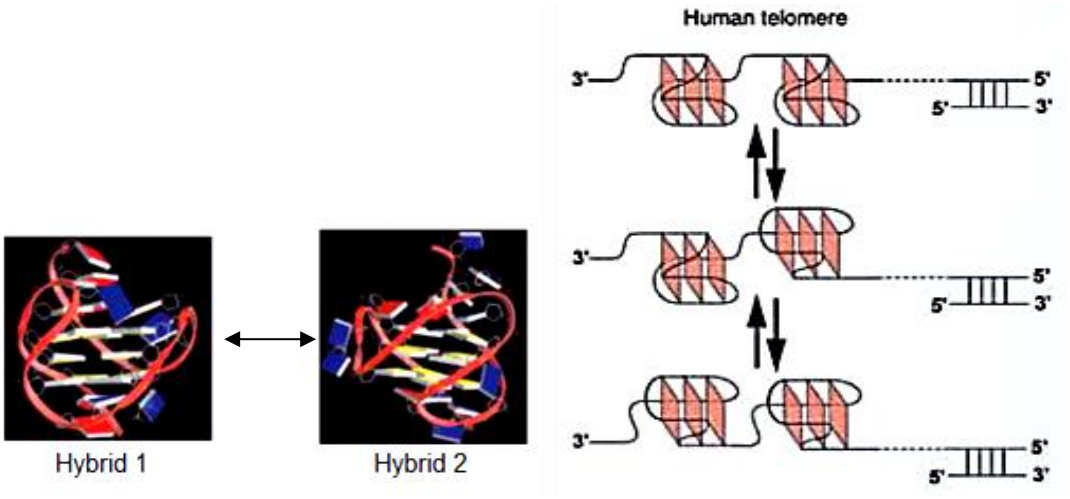


Figure 1.27: (a) Distinct structures of hybrid-1 and hybrid-2, established by NMR studies of telomeric G-quadruplexes (here, guanine, adenine and thymine are in yellow, red and blue blocks, respectively). (b) DNA secondary structure model formed due to compact-stacking of hybrid-type G-quadruplexes in an equilibrium of hybrid-1 and hybrid-2 in K^+ solution.

1.6.3.1 G-quadruplex structures in promoter regions

Quadruplex-forming DNA sequences are not only found in the human telomeric region. They are also observed in various genomic loci and are thought to have a potential role in oncogene regulation.[149, 150] In 2005, two surveys of the human genome by Huppert and co-workers[151] estimated that it contains approximately 375,000 quadruplex-forming sequences, with most located in the promoter regions of genes where they have the potential to regulate gene expression [152-154]. Some major examples of DNA sequences identified as forming quadruplex structures in the human genome are listed in the following table:

Table 1.2: Examples of DNA sequences identified as forming quadruplex structures in the human genome.[154]

Gene	*Sequence*
sc-Myc	
Pu27	TTATGGGGAGGGTGGGGAGGGTGGGGAAGG
Myc-2345	TGAGGGTGGGGAGGGTGGGGAA
Myc-1245	TGGGGAGGGTTTTTAGGGTGGGGA
Myc-22	TGAGGGTGGGTAGGGTGGGTAA
Pu24I	TGAGGGTGGUGAGGGTGGGGAAGG
c-Kit	
c-Kit1	CAGAGGGAGGGCGCTGGGAGGAGGGGCTG
c-Kit2	CCCCGGGCGGGCGCGAGGGGAGGGGAGGC
VEGF	CCCGGGGCGGGCCGGGGGCGGGGTCCCGGCGGGGCGGAG
HIF-1α	GCGAGGGCGGGGGAGAGGGGAGGGGCGCG
Bcl-2	GTCGGGGCGAGGGCGGGGGAAGGAGGGCGCGGGCGGGGA
K-Ras	GGGAGGGAGGGAAGGAGGGAGGGAGGGA
Rb	CGGGGGGTTTTGGGCGGC

1.7 Telomerase and Cancer

In 1997, Shay and co-workers reviewed all possible types of human cancers and established a very strong correlation between telomerase activity and malignancy. This has led the enzyme to become one of the most important tumour markers.[81]

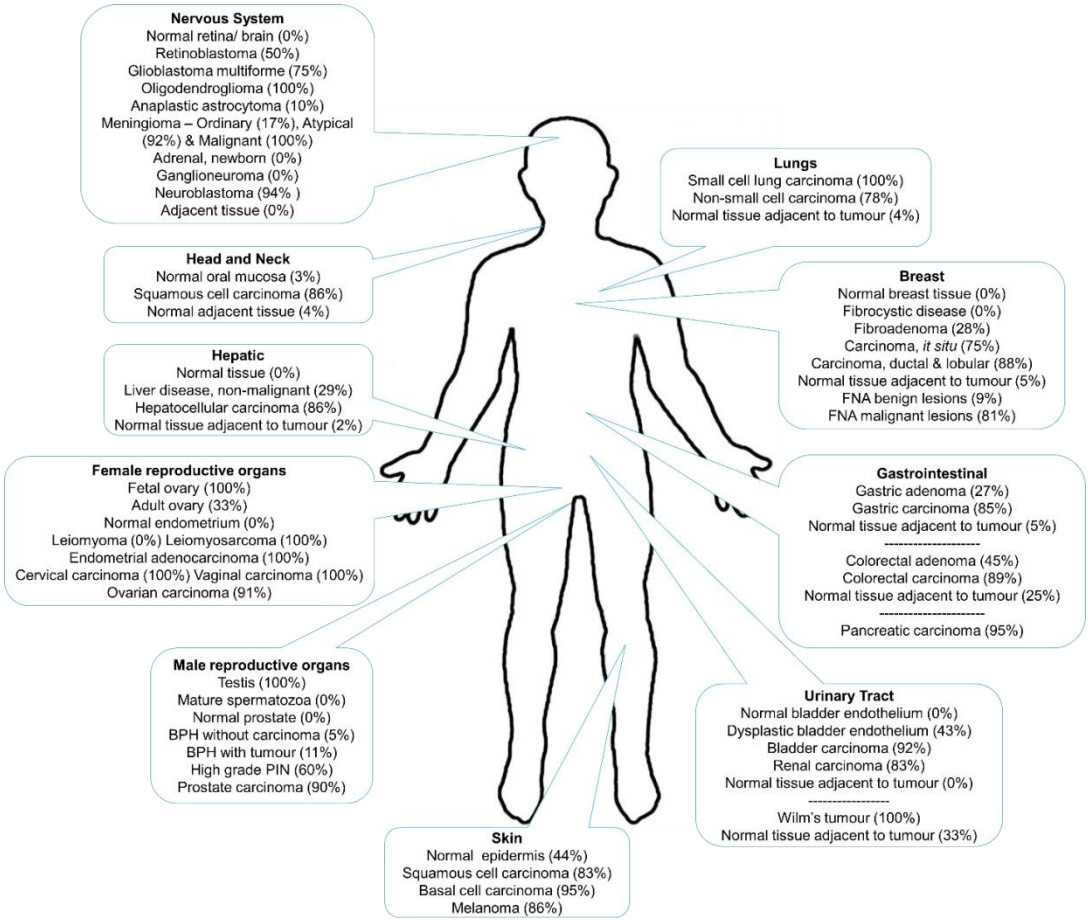


Figure 1.28: Summary of telomerase activity in humans, where tumour samples were evaluated in a TRAP assay (telomerase positive samples are compared to matched control tissue and denoted in %).[81]

1.7.1 Mechanisms of Telomerase Up-Regulation in Cancer

Figure 1.29 describes the detailed mechanisms involved in the up-regulation of telomerase in cancer. Many studies have focussed on the role of transcriptional control of hTERT expression or on phosphorylation-induced increases in hTERT activity.[155-159] Positively acting transcription factors and co-regulators include the Myc-oncogene, the E6 protein of oncogenic human papillomaviruses, NF-κB and β- catenin. Upregulated expression may also occur through loss of pRb, WT1, Menin and TGFβ pathway tumour suppressor function. Protein kinase Cα and PKB can phosphorylate hTERT. Mutations within the core promoter region of hTERT are common, and these may increase hTERT transcriptional activity by creating new consensus binding motifs for ETS transcription factors. hTERT promoter mutations are most common in

malignancies like bladder carcinoma, liposarcomas (*e.g.*, fat cells of deep soft tissue), hepatocellular carcinoma, squamous cell carcinoma (*e.g.*, tongue), medulloblastomas (*i.e.*, malignant primary brain tumour) and glioblastoma (aggressive malignant primary brain tumour).[160-164] Another mechanism of up-regulation is translocation of the hTERT gene to an immunoglobulin gene or other loci in B-cell neoplasms, which presumably contributes to the increased hTERT transcription and telomerase activity observed in these tumours.[165] The human telomerase catalytic subunit (hTERT) and the RNA template molecule (hTR) are both present in very low amount in normal cells, and it is likely that the expression of both must be increased in order to express sufficient telomerase to prevent telomere shortening in cancer cells.[166] PAX8, a protein encoded by the *PAX8* gene, is able to co-ordinately upregulate both hTERT and hTR expression, and the PAX8 expression level correlates with telomerase activity in gliomas.[167] The copy number of the hTERT and hTR genes is commonly increased in human cancers.[168]

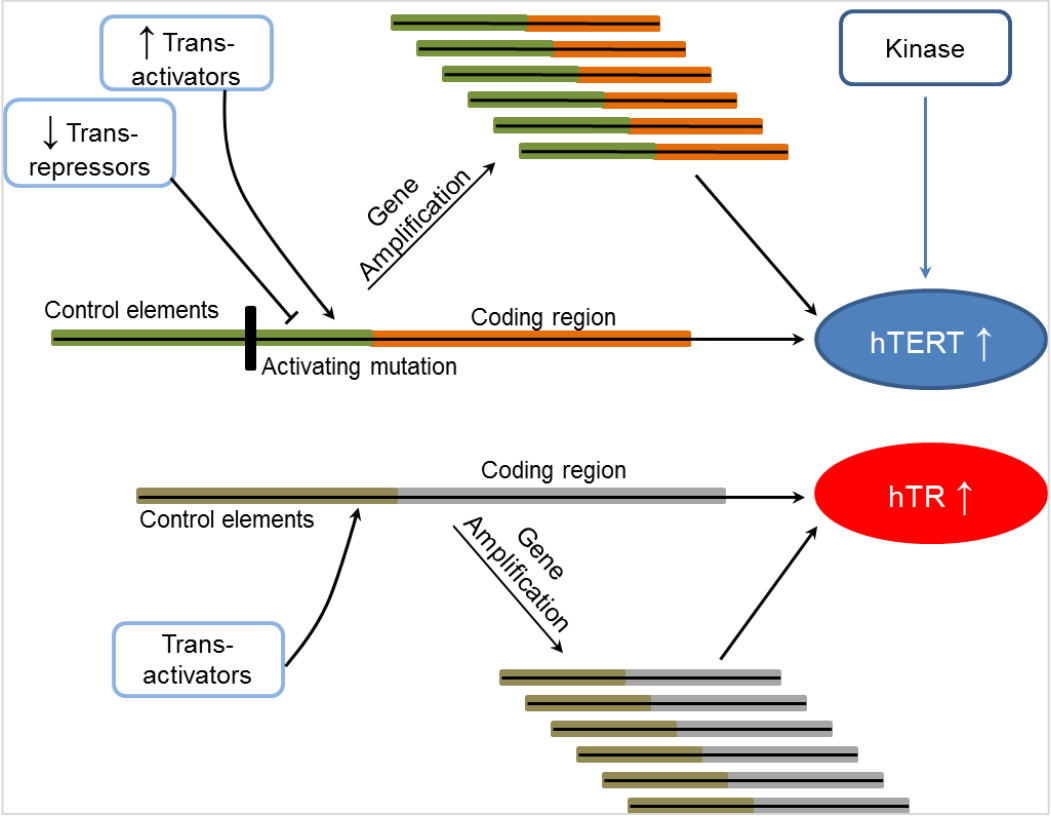


Figure 1.29: Telomerase up-regulation in cancer cells can occur by a variety of mechanisms.

1.7.2 Telomere Lengthening by ALT

Some immortalized cell lines maintain telomere length for hundreds of population doublings in the absence of telomerase activity, and it was therefore deduced that they must have an alternative lengthening of telomeres (ALT) mechanism.[169] In contrast to telomerase activity, where an RNA template is used for the synthesis of new telomeric DNA[170], the ALT mechanism involves the elongation of the telomeric DNA sequence from a DNA template *via* a homologous recombination protein (**Figure 1.30**).[171] Somatic cell hybridization analysis has demonstrated that normal telomerase-negative human cells, as well as telomerase-positive immortalized cells, contain repressors of ALT activity.[169, 172, 173]

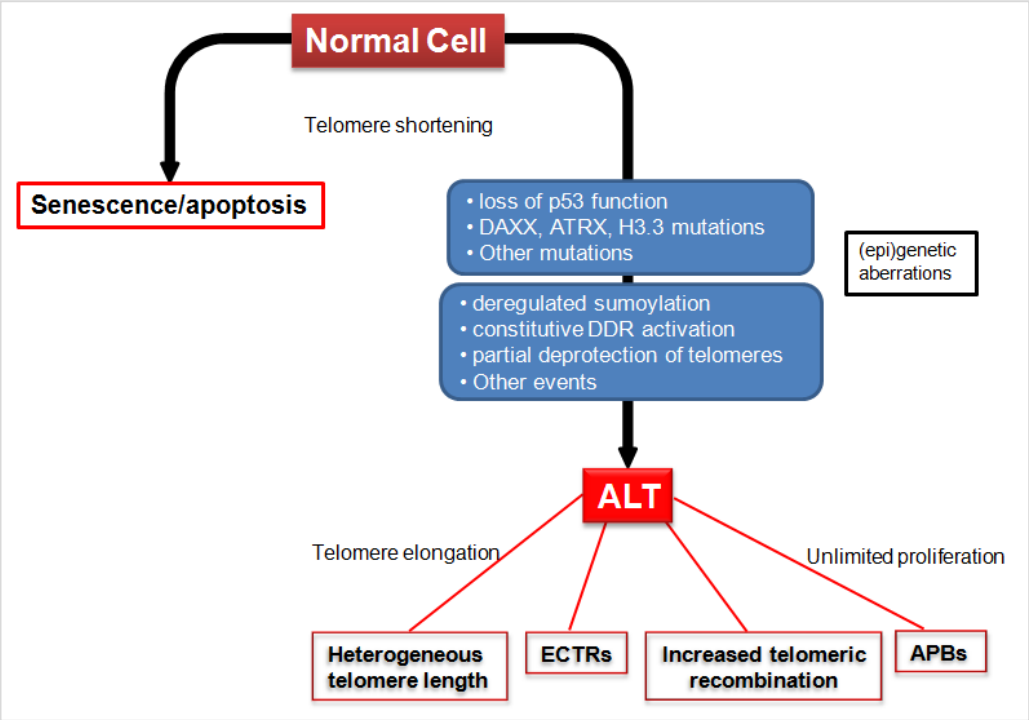


Figure 1.30: Multiple deregulation steps associated progression of the ALT mechanism.[174]

1.8 Telomerase and the Aging Process

Telomeres are considered an important part of the cell fate and aging process as the regulation of cellular responses to stress and growth stimulation, on the basis of earlier cell divisions and DNA damage, are controlled from the telomeric end of a DNA strand. As a minimum, a few hundred nucleotides in the telomere repeats protect each chromosome end to evade the

activation of DNA restoration pathways. In most somatic cells, repair of critically short or 'uncapped' telomeres by telomerase or recombination is limited, which causes apoptosis or cellular senescence due to accumulation of numerous 'uncapped' telomeres and the increased chance of reduced telomere length. Moreover, telomere length is varied in somatic cells and gradually declines with age, which places a block on tumour growth and causes the loss of cells with age.

Aging can be described as the inevitable decay of tissue functionalities that eventually results in death. Such functional degeneration can be a consequence of the loss or weakened function of post-mitotic cells or from failure in substituting similar types of cells due to the functional inability of stem cells to keep up the cellular replication and cell division processes. Variation in DNA repair pathways between germ-line cells and stem cells and differences in somatic cells' size and life expectancy among animals present difficulties in understanding an overview of the molecular mechanism of aging for different species.

It has been reported that p53 plays a comprehensive role in typical growth and tumour formation, life expectancy and overall fitness.[175] In practice, DNA damage signals originate from short telomeres and contribute to the activation of p53 and subsequent cellular responses to stress. In addition, TRF2 binds to the ataxia telangiectasia mutated (ATM) kinase and can inhibit ATM-functionalities, though the DNA damage signals seem to initiate from telomeres during each replication cycle. Thus it is suggested that telomeres can shift between closed and open states (**Figure 1.31**) where the possibility of the open state is proportional to the overall telomere length of the repeat strand. As telomere length gets shorter with age, the number of DNA damage signals originating from short telomeres also increases, subsequently increasing the level of activated p53 and causing either the onset of senescence or apoptosis activation. It has already been confirmed that telomeres are directly responsible for sustained DNA damage signals in senescent cells.[176]

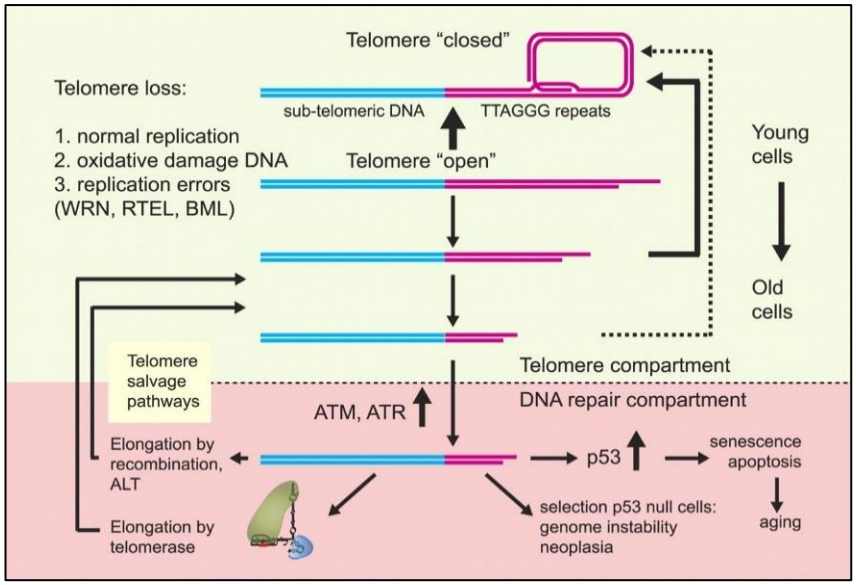


Figure 1.31: Scheme of factors affecting the telomere length in human tissues that also contribute to the aging process.[177]

1.9 Opportunities for Telomerase Targeted Therapies

Complex processes, like accumulation of mutations or alterations in gene expression, lead normal cells to become tumorigenic, where telomerase activation is observed during one of these critical stages.[178] In addition, during malignancy, when cancer cells lose the essential aspects of proliferative control, cells utilize stabilized genetic material afforded by active telomere synthesis to bypass the growth regulatory signals imposed on ordinary tissues. There are many aspects of telomere maintenance in cancers that present potential opportunities for developing targeted therapies. The most obvious of these are the catalytic events involved in synthesis of new telomeric DNA by telomerase or ALT.

The telomerase-catalysed telomere lengthening reaction proceeds via multiple steps, which include binding to its substrate, reverse transcription of telomeric DNA, and translocation of the enzyme on its extended substrate to continue the lengthening reaction. All of these steps could potentially be inhibited (**Figure 1.32**), and inhibitors of the reverse transcription step could target either the template or the catalytic subunit. Telomerase-mediated lengthening

could potentially be inhibited by targeting other steps in the overall process, starting with expression of the subunits and their assembly into the active enzyme complex.

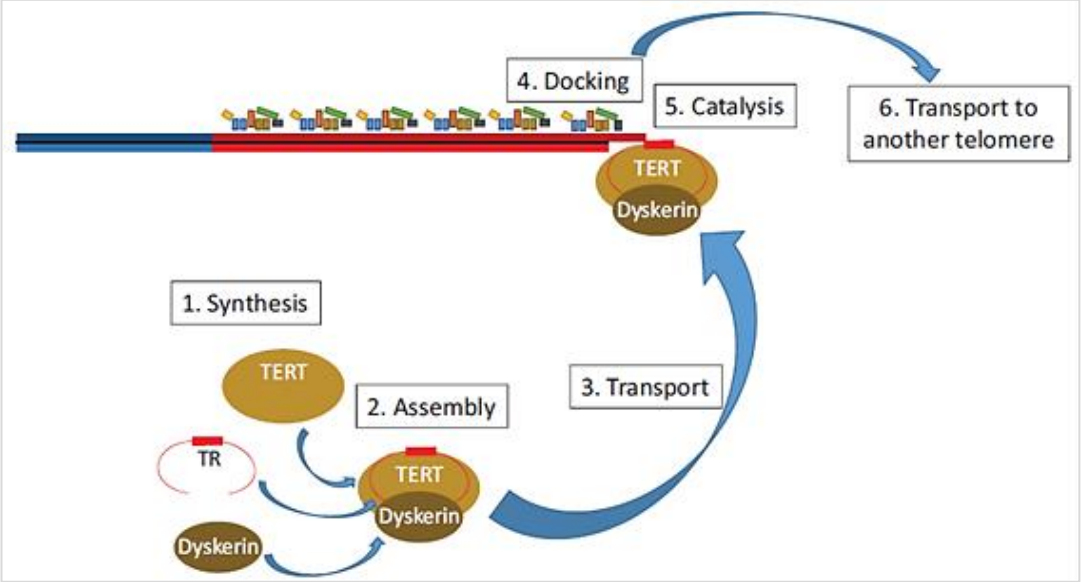


Figure 1.32: The action of telomerase may be inhibited at multiple points in its "life cycle".[83]

Telomeres are particular DNA-protein complexes at terminals of chromosomes that represent a potential target for anticancer agents.[179] Anticancer approaches directed at telomerase inhibition can be varied and possible methods include stabilizing telomeric G-quadruplex structures, RNA interference (RNA) of the hTERT catalytic subunit and stabilizing DNA/RNA hybrid duplex structures towards the large protein surface of the telomerase enzyme. The steps in the ALT process involve homologous recombination (HR) proteins and DNA synthesis, which are processes vital for the repair and replication of DNA in normal cells. Although there may be a therapeutic window where cancer cells, especially those which have lost p53 function, are more vulnerable to inhibition of these functions than normal cells, the likelihood of toxic side-effects of such inhibition is high.

Peptides generated by degradation of hTERT are presented on the cell surface, and cytotoxic T-lymphocytes (CTL) can recognise these peptides and kill the cell. This makes telomerase a potential target for the development of immunotherapies.[180, 181]

1.9.1 Telomeric DNA/RNA Hybrid Duplex as a Target in Cancer

It is possible to design drugs to bind to specific DNA structures, and the G-rich nature of the telomeric sequence has led to extensive efforts to develop drugs that target the telomeric G-quadruplex DNA. The DNA/RNA hybrid duplex was discovered in 1960. Inhibiting telomerase by binding to the hybrid duplex is a new approach for cancer treatment. This duplex structure usually forms during the catalytic cycle of telomerase activity. It forms during the original annealing of the substrate telomere to the template RNA and during its extension.[107, 182]

Recently, it has been established that telomere uncapping results in rapid cell death. Uncapping of chromosomal DNA can take place due to 'damage' signals at the G-strand overhang and/or the corresponding lagging strand that can lead to rapid apoptosis or genomic instability. Thus, ligands that 'uncap' and/or directly target telomeres are likely to be more efficient as a monotherapy and can act more rapidly compared with enzyme inhibitors alone.

The telomeric DNA/RNA hybrid (DRH) duplex is a novel therapeutic target for potential anticancer agents that would work through the inhibition of telomerase. Such hybrid duplex structures usually form only during the telomerase catalytic cycle (though such hybrids can be formed during the transcription process, which are unstable and dissociate during the translation process), thus making them highly selective targets for rational drug discovery.[107, 182] In addition, DRH duplexes are less common structures compared with DNA/DNA duplexes in cells[183], and contain the specific sequence of TTAGGG. Furthermore, as the DRH duplex is bound tightly by the protein unit of telomerase, ligands that bind to this duplex should block access to the unique protein surfaces of telomerase. Synthesis of the telomeric repeat involves the 3′-end of the DNA primer base pairing with the corresponding RNA template to form a DRH duplex positioned within the active site. DNA polymerization then continues through RNA template reverse transcription to add the telomeric repeat (TTAGGG) onto the 3′-end of the DNA primer. Thus the template is regenerated through a multi-step process that involves separation of the DRH, and realignment of the RNA template relating to the DNA primer, followed by re-formation of the DRH.

Moreover, diversity in hybrid structures (due to the presence of different nucleobases, purine and/or pyrimidine in the DNA strand) offer different backbone conformations and groove geometries, usually the A-form, in the DRH duplex that are distinctive features against which to target selectively small molecule analogues.[184]

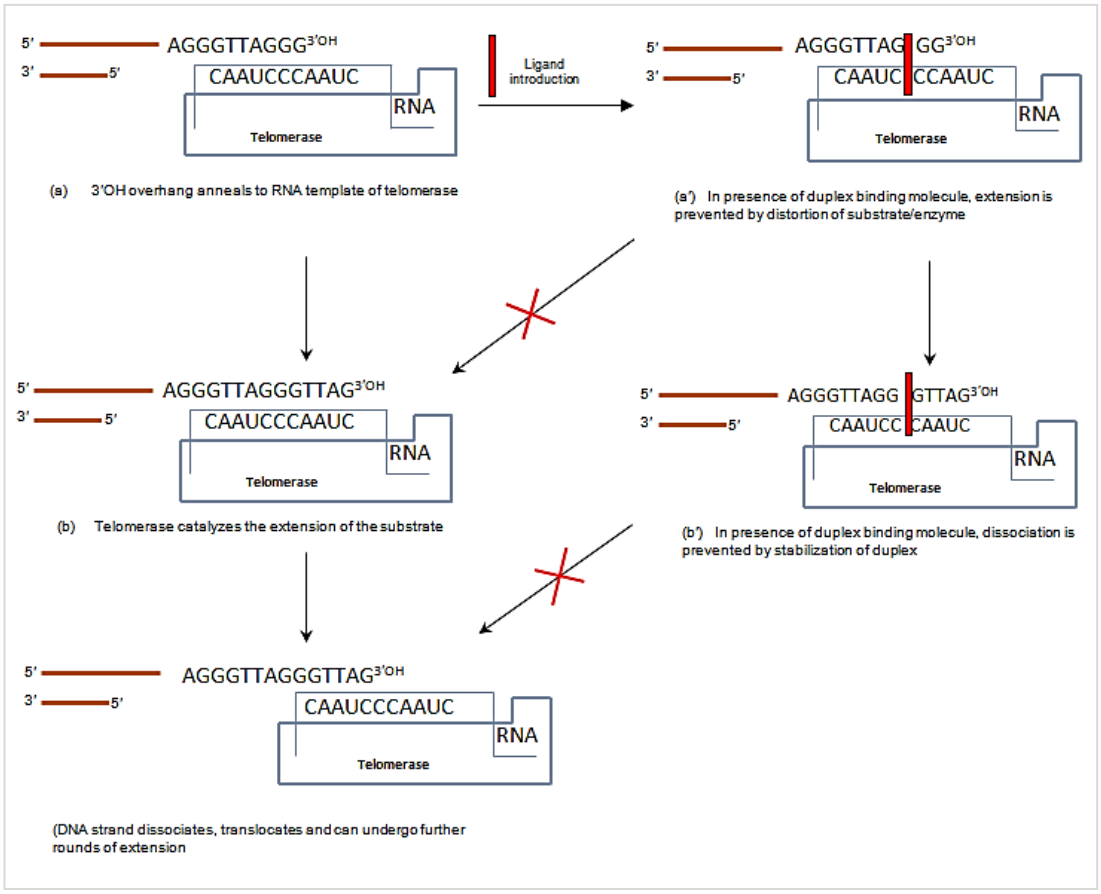
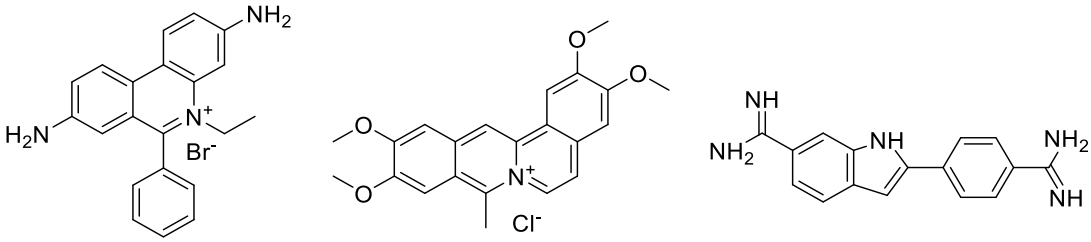


Figure 1.33: Possible ligand binding mechanisms to inhibit telomerase activity by stabilizing the DNA/RNA hybrid duplex, as proposed by West and co-workers.[182]

As the telomeric DNA/RNA hybrid duplex is a new anticancer drug target, there are only a limited number of reported molecules that can selectively stabilize the DRH duplex and stop telomere elongation in cancer cells. However, other molecules have previously been reported that target other DNA/RNA hybrid duplexes formed within the cells. In 1999, Ren and co-workers reported 15 different DNA binding agents and explored them by using a thermodynamically rigorous competition dialysis procedure, where single-stranded, a variety

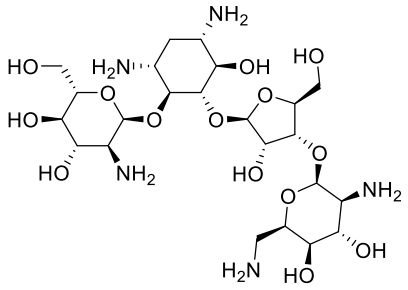
of double-stranded, multi-stranded triplex and tetraplex forms of DNA, left-handed Z-DNA, RNA, and a DNA/RNA hybrid duplex (other than the telomeric DNA/RNA hybrid duplex) were used.[185] The authors analysed daunorubicin, ethidium bromide, coralyne chloride, chromomycin, distamycin, actinomycin D, berberine, berenil, BePI, 1-pyrenemethylamine hydrochloride, 5,10,15,20-tetrakis[4-(trimethylammonio)- phenyl]-21*H*,23*H*-porphine, DAPI, H2TMPyP [meso-tetrakis(N-methyl-4-pyridyl)porphine, NMM (N-methyl mesoporphyrin IX), and netropsin in their study. Among fifteen molecules, ethidium bromide (**1.17**), coralyne chloride (**1.18**) and DAPI (4′,6-diamidino-2-phenylindole) (**1.19**) showed binding affinity for poly(rA):poly(dT) hybrid duplexes, which was further confirmed by reviewing equilibrium binding assays in previous literature. However, daunorubicin, an anticancer agent, did not show any affinity for the hybrid duplex.



Ethidium bromide (**1.17**) Coralyne chloride (**1.18**) DAPI (**1.19**)

Figure 1.34: Some reported molecules that showed binding affinity towards a DNA/RNA hybrid duplex (poly(rA):poly(dT)) sequence.[185]

Structure-selective and sequence-selective ligands for DNA/RNA hybrids have a diverse array of possible pharmaceutical applications. For example, the first-formed product of HIV reverse transcriptase was a DNA/RNA hybrid duplex.[186] Telomerase is a specialized reverse transcriptase enzyme, which has protein structure domains closely-related to the HIV reverse transcriptase enzyme.[187] Besides, since the DNA/RNA hybrid duplex's thermodynamic stability regulates telomerase activity, this can be a small-molecule targeting strategy.[107, 188, 189] Barnieri and co-workers also observed that the binding of paromomycin (**1.20**) to the DNA/RNA hybrid duplex inhibits both RNase H- and RNase A-mediated cleavage of the RNA strand.[190]



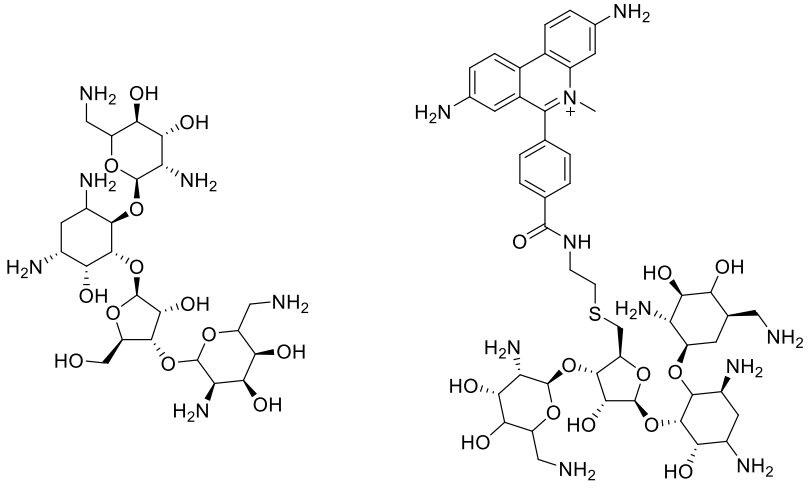
(**1.20**)

Figure 1.35: Structure of Paromomycin.

In 2007, Rangarajan and co-workers made some phenanthridine derivatives, which included a 5-hexanoic acid derivative, a 5-benzylic acid derivative, *N*-glutaryl derivatives and the carboxamides of all of these.[188]

According to their report, only one compound (*i.e.*, 3,8-diamino-5-(4-carbamoylbenzyl)-6-phenyl-1,2-dihydrophenanthridin-5-ium, $^{Telo}IC_{50}$ = 2.9 μM) gave similar telomerase inhibition of standard ethidium bromide ($^{Telo}IC_{50}$ = 2.7 μM); however, the other derivatives showed much higher inhibitory concentrations, ranging from 5.6 μM to >80 μM.

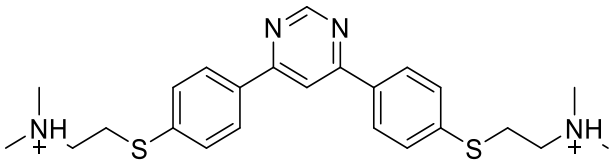
In 2008, Shaw and co-workers reported a neomycin-ethidium chloride conjugate that binds to the DNA/RNA hybrid duplex (poly(dA):poly(rU)) and shows a greater binding constant (K_a = 4.8±1x10^6) towards the hybrid sequence than its parent molecules, ethidium bromide and neomycin (**Figure 1.36**).[191]



(**1.21**) (**1.22**)

Figure 1.36: Structures of neomycin (**1.21**) and neomycin-methidium conjugate (**1.22**).

In 2010, Wheelhouse and co-workers made a significant advance in DNA/RNA hybrid duplex research by investigating a molecule that can preferentially bind to telomeric quadruplex and triplex DNA structures, as well as to the hybrid duplex poly(dA)·poly(rU).[192] Competition dialysis showed a 20-fold preference for duplex DNA associated with the poly(dA)-poly(rU) hybrid duplex compared with equivalent RNA, a 3-fold preference over duplex DNA, and 7-fold preference over the alternative poly(rA)-poly(dT) hybrid sequence.



(**1.23**)

Figure 1.37: Structure of a novel diaryl substituent and pyrimidine analogue with a selective affinity for DNA/RNA hybrid duplexes.[192]

Upon reviewing all available literature, there are no reported molecules that target the telomeric DNA/RNA hybrid duplex that forms during the catalytic stage of telomerase expression.

1.9.2 Telomeric G-quadruplex as a target

G-quadruplexes (G-tetrads or G_4-DNA) are formed from nucleic acid sequences that are rich in guanine and so have the ability to form four-stranded structures. Four guanine bases can join together through Hoogsteen hydrogen bonding to form a square planar structure, which is known as a guanine tetrad, and two or more guanine tetrads can stack on top of each other to form a G-quadruplex structure. The G-quadruplex is stabilized by the presence of a cation, especially potassium (K^+), which sits in the middle channel between each pair of tetrads.[193]

Base stacking interactions, hydrogen bonds, electrostatic interactions and the hydration shell are the key factors for stabilizing the quadruplex structure.[146] Guanine bases are the key requirement for the formation of the base stacking within the G-quadruplex structure. However, due to the arrangement of the central guanine O6 carbonyl groups of the G-quartet, which

form a square planar structure for each tetrad giving a twist of 30° and rise of 2.73 Å between each tetrad step to generate a bi-pyramidal antiprismatic arrangement for the eight O6 atoms (**Figure 1.38**), quadruplex structures become unstable in the presence of stabilizing factors. This arrangement of O6 carbonyl atoms forms negatively charged cavities between the G-tetrads which cations can neutralise. The overall stability of the folded quadruplex largely depends on the selection of a suitable cation and its size and charge.

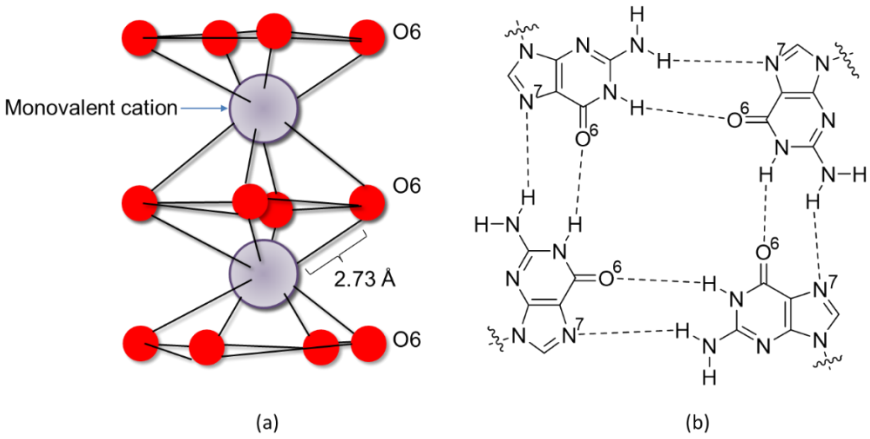


Figure 1.38: (a) Metal ion is shown coordinated between eight carbonyl O6 atoms with an average cation-oxygen distance of 2.73 Å; (b) structural illustration of O6-NH bonding for each quartet (*note, hydrogen bonds are not of an appropriate length*).

In 1991, Zahler and co-workers demonstrated that potassium ions (K^+) stabilize G-quadruplex structures and are ultimately capable of inhibiting telomerase activity.[135] Since then, G-quadruplex structures were targeted to inhibit telomerase activity. Some research groups working in this field are searching for small organic molecules that can bind to stabilize G-quadruplexes and ultimately block telomerase enzyme activity.

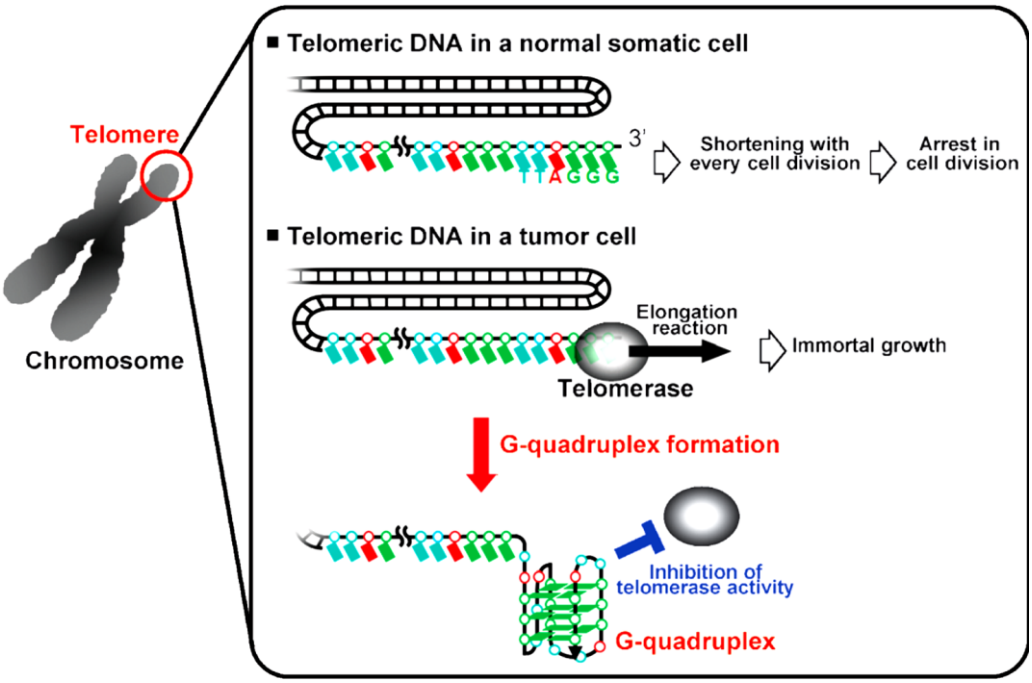


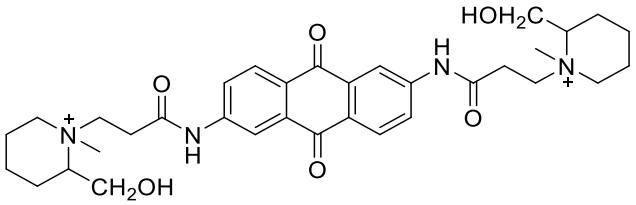
Figure 1.39: Roles of telomeric DNA and telomerase; G-quadruplex formation can inhibit telomerase activity.[194]

In 2011, Alcaro and co-workers reported that the main six diverse conformations of the unimolecular hTelo sequence in the presence of Na^+ and K^+ ions represent unique targets against which to design small molecules with improved specificity. The six conformations can be found in the Protein Data Bank (PDB)[195] under the codes 143D (antiparallel, in Na^+)[144], 1KF1 (parallel, in K^+)[196], 2HY9 (hybrid-1, in K^+)[197], 2JPZ (hybrid-2, in K^+)[198], 2JSL (hybrid-1, in K^+)[145], and 2JSM (hybrid-2, in K^+)[145].

A number of molecules have been studied to date, which can stabilize the G-quadruplex structure either by end stacking or groove binding. Mostly the G-quadruplex stabilization happens by stacking and electrostatic interaction. Thus G-quadruplex binding ligands are normally planar aromatic molecules that are prone to stack with G-tetrads, while some of them are positively charged or have hydrophilic groups to favour electrostatic interaction. It has also been proposed that *bis*- and *tris*-intercalating drugs show promising activity and selectivity.[199, 200]

In 1997, Sun and co-workers first reported a 2,6-diamidoanthraquinone analogue as a human telomerase inhibitor.[201] In their study, they demonstrated the G-quadruplex inhibition

mechanism for this anthraquinone-type molecule by ^{1}H NMR. The group developed a series of 2,6-diamidoanthraquinones as DNA-interactive agents with reduced affinity for duplex DNA. A standard primer extension assay (*i.e.*, TRAP assay) was employed to study the relative inhibitory activity of **1.24** towards human telomerase, with an IC_{50} value of 23 μM and complete inhibition of telomerase activity at 100 μM.



(1.24)

Figure 1.40: 2,6-Diamidoanthraquinone analogue.

Following Sun's work, several series of compounds have been developed and explored by different research groups as a G-quadruplex ligand and most of them were developed to reduce toxicity, and improve selectivity and efficacy. After observing the G-quadruplex targeting potential of anthraquinone type molecules, a number of compounds with similar structural features (*e.g.*, di-substituted tricyclic central core) like acridines and fluorenones were developed, where acridines showed more potent telomerase inhibitoion with lower toxicity. In 1999, Perry and co-workers first described fluorenone-type compounds in comparison with the anthraquinone-type compounds.[202] This fluorenone series of molecules was reported as a wide range of telomerase inhibitors with potential activity (8-12 μM) against telomeric G-quadruplex sequences. The authors reported **1.25**, with 1-(2-hydroxymethyl)piperidinyl terminal side arms, as the most potent telomerase inhibitor with an IC_{50} value of 8 μM. BRACO19 (**1.26**), a 3,6-*bis*(aminoalkyl) substituted acridine, is a rationally designed tri-substituted acridine derivative that appears to directly target telomeres. It is a member of a series of potent inhibitors that were designed by computer modelling to exploit the unique structural features of the DNA quadruplex structure. Specifically, this molecule was designed to interact with the three minor grooves of the G-quadruplex in addition to stacking on the terminal G-quartet. BRACO19 is a potent G-quadruplex stabilizing agent with a high

ΔT_m of 27.5 °C at 1 μM concentration. Additionally, it has moderate inhibitory activity in the TRAP assay (*i.e.*, IC_{50} = 0.09 μM).[203-205] In addition, BRACO19 has potent inhibitory activity against a number of tumour cell lines [206] and has been shown to inhibit the catalytic function of telomerase in human cancer cells and destabilize the telomere capping complex [207]. However, lack of permeability and a small therapeutic window limited the use of BRACO19 as a chemotherapeutic agent.

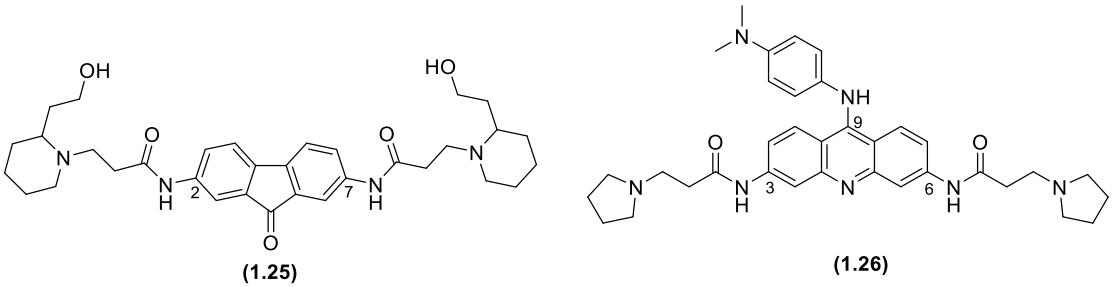


Figure 1.41: Example of fluorenone and acridine molecules.

Riou and co-workers identified a series of triazine derivatives in 2002 with potent anti-telomerase activity at the nanomolar level, where compounds **1.27** and **1.28** gave ΔT_m values of 20 °C and 8 °C in a FRET melting assay and IC_{50} values of 0.041 μM and 0.13 μM in the TRAP assay, respectively.[208] Molecular modelling studies were used in 2010 by Sparapani and co-workers to design a series of di-substituted triazole-linked acridine compounds.[209] Among the library of novel compounds, **1.29a** and **1.29b** had good selectively toward the hTelo quadruplex sequence compared with duplex DNA; at a 1 μM concentration they stabilized the hTelo G-quadruplex by 15.8 and 15.5 °C, respectively, relative to 0.4-0.6 °C stabilization for ds-DNA. Non-specific quadruplex binding ability limits the exploration of this type of molecules.

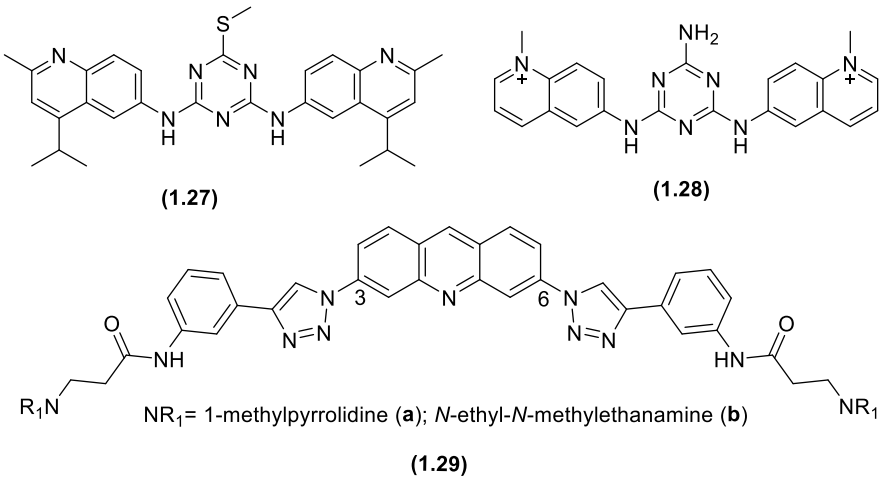
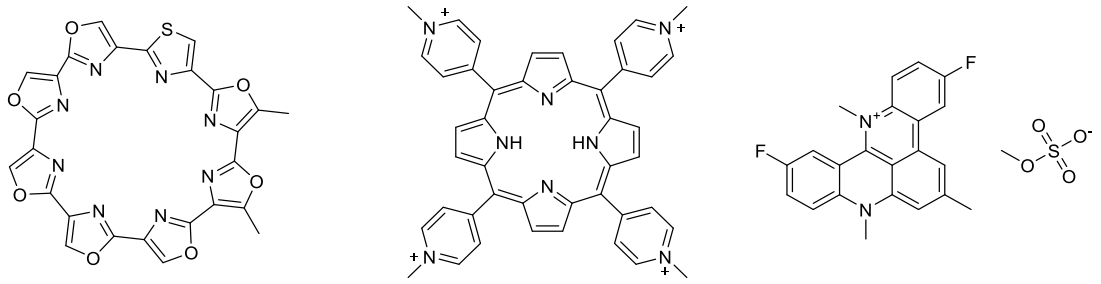


Figure 1.42: Example of triazine type molecules.

Telomestatin (**1.30**) is a macrocyclic compound first isolated from *Streptomyces anulatus*[210] and later synthesized in the laboratory as *R*-telomestatin[211]. It is one of the most effective small-molecule inhibitors of telomerase. With a flat and planar structure and a resemblance to the structure of the G-quartet, telomestatin was recognized as a G-quadruplex binding agent, whichwas successively confirmed by molecular modelling studies. Telomestatin showed preferential interaction with intramolecular compared with intermolecular G-quadruplex structures with a 70-fold selectivity for intramolecular quadruplex structures over the duplex DNA. Telomestatin demonstrated favourable anticancer activity against different human cancer cell lines, where treatment with telomestatin for 3-5 weeks gave repressed telomerase activity with reduced telomere length and apoptosis.[212] Telomestatin has been studied extensively to define the mode of action and it was observed that telomestatin leads to the telomere dysfunction and impair chromosomal integrity. In 2003, Tauchi and co-workers reported that telomestatin can activate ATM and Chk2 in the BCR-ABL positive leukemia cell line and later can raise the expression of p21 and p27, which was thought to be a possible way to induce an ATM-dependent DNA damage response.[213] In 2006, inhibition of DNA binding of POT1 and TRF2 and the loss of G-tails were also reported as possible quadruplex stabilization activity by telomestatin.[214, 215]



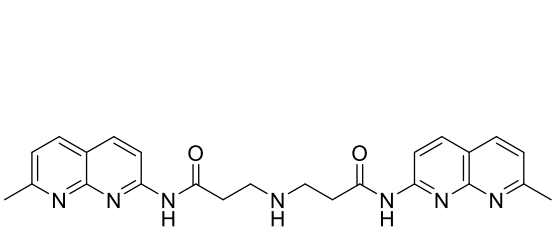
(1.30) (1.31) (1.32)

Figure 1.43: Example of cyclic porphyrines and five-ring acridine.

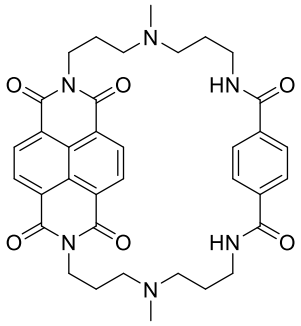
TMPyP4 (**1.31**) is a molecule which was considered by using structure-based drug-design. The molecule was selected as a G-quadruplex binding agent based on its physical characteristics like the presence of a fused planar ring structure, a positive charge at four termini, and suitable size to stack with the G-tetrads. TMPyP4 showed substantial selectivity for the G-quadruplex structure over the duplex DNA, and could efficiently inhibit human telomerase in HeLa cell extracts. Both *in vivo* and *in vitro* results showed that TMPyP4 can decrease the rate of proliferation of sea urchin embryos and block the cells at the mitosis stage. Moreover, it causes the formation of anaphase bridges and demonstrated an antiproliferative effect due to chromosome destabilization.[216] In 2002, it was reported that TMPyP4 interacts with the c-Myc quadruplex structure of the promoter region and thus can downregulate c-Myc expression. It has also been postulated that TMPyP4 contributes to the lowering of hTERT expression. [217] However, a major drawback of TMPyP4 to be a G-quadruplex target agent is its selectivity as TMPyP4 has the ability to bind to duplex DNA, DNA/RNA hybrids and triplex DNA.[218] The five-ring acridine, RHPS4 (**1.32**) was originally designed to inhibit telomerase due to structural resemblances to previously identified G-quadruplex structure binding telomerase inhibitors (*e.g.*, BRACO19).[219] Long-term exposure (*i.e.*, 15 days) of non-acute cytotoxic concentrations of RHPS4 with 21NT breast cancer and A431 human vulval cancer cells demonstrated a decrease in cell growth and a noticeable decline in cellular telomerase activity. Moreover, RHPS4 was also reported as a quadruplex interactive agent as it can stack at the end of G-tetrad.[220]

In 2003, Nakatani and co-workers described the binding of a series of naphthyridine dimers (*e.g.*, **1.33**) to the hTelo sequence[221], where in a preliminary FRET assay using a ligand

concentration of 30 μM in Na-cacodylate buffer, compound **1.33** gave ΔT_m values of 45.6 °C and 63.6 °C for the secondary structures of the 9-mer d(TTAGGGTTA) and 15-mer d(TTAGGGTTAGGGTTA), respectively. Cyclic naphthalene diimide (**1.34**) has also been reported as a telomeric quadruplex agent, binding with an IC_{50} value of 0.9 μM in the TRAP assay.[222]



(1.33)



(1.34)

Figure 1.44: Examples of naphthyridine dimer and cyclic naphthalene structures.

In different studies, cryptolepine-based (*i.e.*, fused indolo-quinoline structure) compounds (**1.35**)[223], phenanthroline-based molecules (**1.36**)[224], berberine analogues (**1.37**) and its dimer (**1.38**)[225, 226] , isaindigotone derivatives (**1.39**)[227] were also reported as telomeric quadruplex interacting agents. Several attempts were also made by different research groups to identify naturally occuring small molecule (*e.g.*, thymoquinone[228], Cristatin A[229]) . In 2014, Dhamodharan and co-workers reported 1,8-naphthyridine and 1,10-phenanthrolinein based *bis*-benzimidazole carboxamide analogues (**1.39** and **1.40**) as quadruplex-stabilizing agents.[230] Moreover, *bis*-benzimidazole-type compounds are well known for their biological properties.[231-233] 1,3-Phenylene-*bis*(piperazinyl benzimidazole) skeleton based compounds (*e.g.*, **1.41**)[234], oligo-oxyethylene (polyglycol) spacer connected *bis*-benzimidazole molecules[235], benzimidazole-carbazole conjugated compounds[236] were also reported to be telomeric G-quadruplex stabilizing agent.

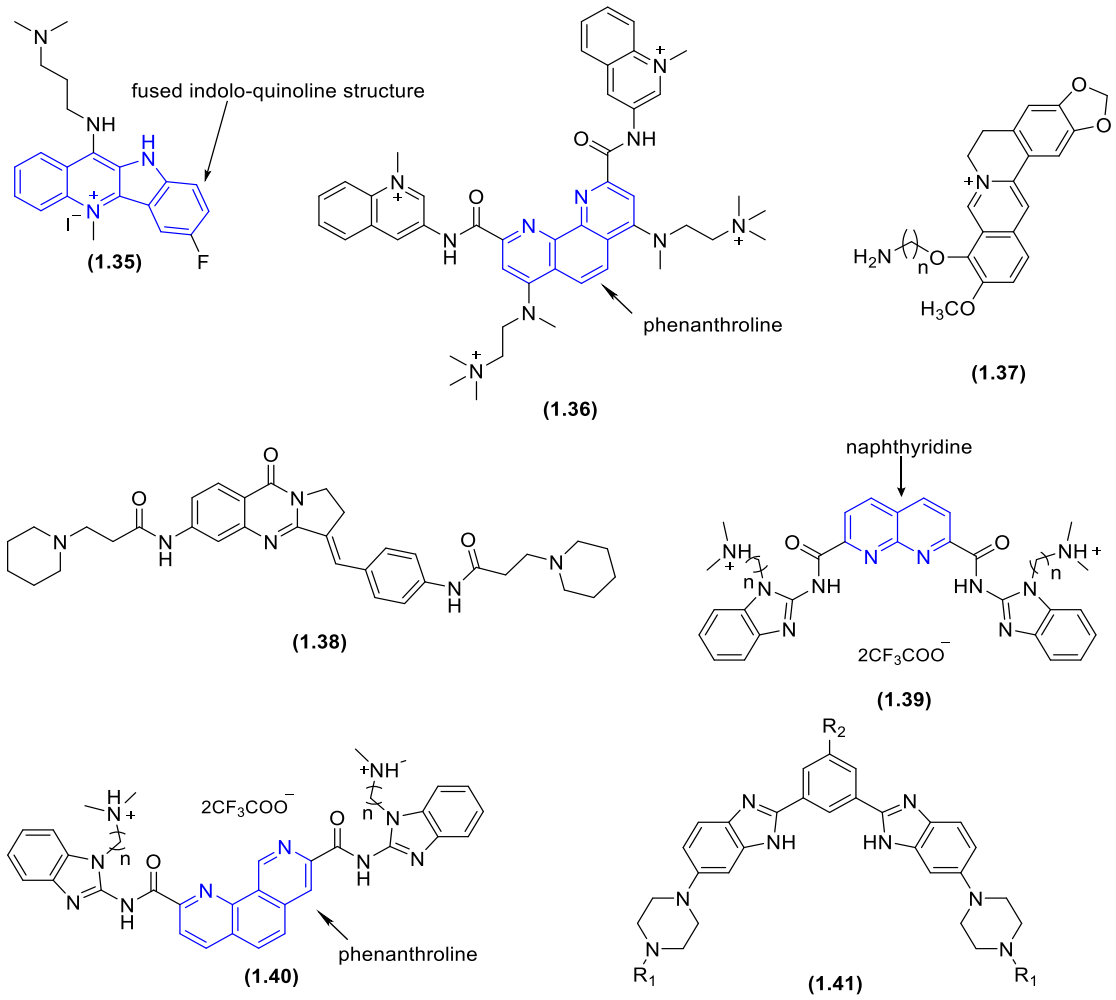


Figure 1.45: Examples of reported non-cyclic aromatic structures with planar shape and a positively charged moiety.

Urea-based alkylaminoanilino analogues (*i.e.*, non-polycyclic aromatic ligands) were also described by Drewe and co-workers in 2008.[237] The result of a FRET melting assay showed that they gave 1.0-14.1 °C and 0.0-5.5 °C stabilization toward F21T and duplex DNA sequence, respectively, at a 1 μM concentration (whereas BRACO19 provided 25.9 °C stabilization for F21T and 11.2 °C for a control duplex DNA sequence). Such findings inspired researchers to look for amide linked small molecules which can meet the pharmacokinetic requirements to become a drug candidate.

The G-quadruplex structure is a DNA secondary structure which has been established as a potential regulatory element in biologically significant regions, for example in human telomeres and in growth related genes' promoter regions. The presence of quadruplex structures in

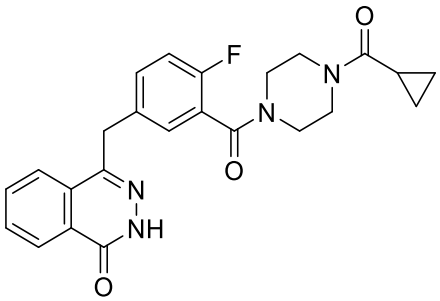
diverse regions of the genome present a design challenge to develop small molecules based on targeted anticancer therapy. For this reason, a number of research groups worldwide are working to solve the problem and identify a suitable drug candidate for a quadruplex based chemotherapy.

1.9.3 Antisense oligonucleotides

Geron Corporation developed an antisense oligonucleotide, Imetelstat (GRN163L), targeting hTR. GRN163L is a lipid-conjugated 13-nucleobase-long oligonucleotide sequence which is complementary to the RNA template of telomerase, and has high binding affinity to the RNA template, thereby directly inhibiting telomerase activity. GRN163L stops telomeric activity due to a direct antagonistic effect on the telomerase RNA template. It should be noted that GRN163L is not a normal antisense oligonucleotide, but with its 5′-Palmitoyl-TAGGGTTAGACAA-3′ oligonucleotide chain partially overlaps the hTR template with highly specific binding affinity at its active site that ultimately causes the complete inhibition of telomerase. GRN163L has demonstrated a favourable pharmacokinetic profile with minimal side effects, and is now in phase 1 and 2 clinical trials.[238]

1.9.4 Targeting telomere and telomerase-associated proteins

Tankyrase inhibitors have implications in the inhibition of telomerase activity and function. Tankyrase-1, a TRF1-interacting tankyrin-related poly(ADP-ribose) polymerase (PARPs), is an enzyme that in humans is encoded by the TNKS gene. Tankyrase-1 releases from the telomere and allows the telomere access to the telomerase. Thus, tankyrase inhibitors can play a vital role in inhibition of telomere functions by not allowing TRF1 to dissociate from the telomere. Olaparib (AZD-2281, **1.42**) is a FDA-approved PARP inhibitor and targeted therapy of cancers with hereditary BRCA1 or BRCA2 mutations, including ovarian, breast, and prostate cancers.[239]



Olaparib (**1.42**)

Figure 1.46: FDA approved poly-ADP ribose polymerase (PARP) inhibitor.

Tankyrase-1 is usually activated by binding to DNA breaks and subsequently plays roles in DNA base excision repair, and is also responsible for modifying the telomere structure by exposing telomeric DNA to telomerase. As TRF1 is a negative regulator of telomere length, removing of TRF1 from telomeres allows telomerase access to telomeric DNA and thus leads to the telomere elongation process in telomerase positive cells. Tankyrase-1 causes poly (ADP-ribosyl)ation of TRF1 protein and reduces TRF1 binding to telomeric DNA, leading to the ubiquination process and ultimate release of TRF1 from the DNA strand. Moreover, tankyrase-1 can directly contribute to telomere elongation by acting as a positive regulator of telomere length and causing the dissociation of TRF1 from telomeric DNA. Therefore, targeting tankyrases with PARP inhibitors is a novel anticancer therapy approach.[240]

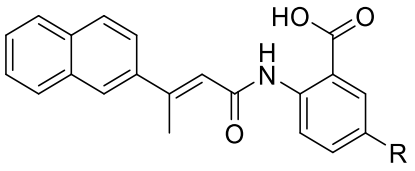
Inhibition of Heat Shock Protein 90 (HSP90) is another telomerase targeted anticancer strategy. HSP90 is a molecular chaperon protein, which is important for protein folding and regulation. In cancer, HSP90 is found to be amplified, which makes HSP90 inhibitor as an appealing research area to develop chemotherapeutic agents. Geldanamycin is a good example of an HSP90 inhibitor. Originally, the drug was developed as a kinase inhibitor but later it was evidenced that geldanamycin can block the ATP-dependent binding of HSP90 to p23, which subsequently disrupts the chaperone assembly causing telomerase inhibition.[241] However, due to concurrent high hepatotoxicity and lower solubility of geldanamycin, two more analogues 17-AAG and 17-DAMG were developed and are in phase-I and phase-II clinical trials at the moment, respectively.

T-oligo is another novel and specifically targeting anticancer agent, which is comprised of an 11-nucleobases long, single stranded, oligonucleotide sequence (GTTAGGGTTAG), and

homologous to the telomeric overhang sequence. Studies have demonstrated that various T-oligo treated various tumour cell lines can undergo active DNA damage responses, facilitated through the ATM kinase and its effector proteins (*i.e.*, p53, p57, p95/Nbs1 and E2F1), which are similar to DNA damage responses originated due to uncapping of telomeres and knockdown of TRF2.[242] Some other studies have shown inhibition of angiogenesis in cancer cell lines due to decrease signal in the VEGF receptor[243] and lowered oxidative damage to cells after treatment of T-oligo[244]. However, T-oligo causes transient cell cycle arrest in normal cells.

1.9.5 Small molecule inhibitors of hTERT

Damm and co-workers reported BIBR1532 (**1.43 a**) and BIBR1591 (**1.43 b**), naphthalene-type molecules, as highly selective telomerase inhibitors. They demonstrated that both compounds inhibited the in *vitro* processivity of telomerase in a dose dependent manner, with BIBR1532 and BIBR1591 giving IC_{50} values of 93 and 470 nm with a nuclear extract from HeLa cells. It was reported that BIBR1532 works by directly inhibiting the telomerase enzyme *via* its active site.[245] It has also been reported that BIBR1532 can competitively bind to the active site of hTERT though it does not block the template copying site for the telomere elongation process. Thus it is hypothesised that the molecule either influences the enzyme-DNA-substrate complex translocation or causes enzyme dissociation from the telomeric end of DNA.



(1.43)

(a) BIBR1532, R=H
(b) BIBR1591, R=morpholin-4-yl

Figure 1.47: Structure of BIBR1532 and its analogue, small molecules targeting hTERT.

1.9.6 Telomerase based immunotherapies

In general, immunotherapy based treatment involves controlling the human immune system to destroy cancer cells. Typically, there are two major antigen identifying molecules present in humans, cytotoxic T-lymphocytes (CTL) and CD4+ helper T-lymphocytes, which are responsible for destroying cells. Within telomerase, hTERT is a potential target antigen and

immunotherapy depending on hTERT is considered to be relatively less toxic as the expression of hTERT is related only with telomerase, which is much lower (almost null) in almost all somatic cells.[180]

Two potential telomerase vaccines have been developed, which are now in phase III clinical trials. GRNVAC1 isolates dendritic cells and the RNA that codes for the telomerase protein and puts them back into the patient to make cytotoxic T-cells that kill the telomerase-active cells. GV1001 (KAEL-GenVax Co Ltd, Korea) is an hTERT-peptide based vaccine and is recognized by the immune system that reacts by killing the telomerase-active cells. VX-001, a cryptic peptide based vaccine, has completed the phase I/II clinical trial with different stages of cancer patients with non-small cell lung cancer, breast cancer, melanoma and cholangiocarcinoma. Due to good tolerability and positive correlation between immune response and clinical responses in patients with non-small cell lung cancer, VX-001 has been scheduled for phase III clinical trial in non-small cell lung cancer patients.

1.10 Telomerase inhibition measurement

Sequential loss of the telomere ends of a chromosome due to the semi-conservative DNA replication mechanism is the key scheduling mechanism that controls the cells doubling in number. On the other hand, the reverse transcriptase enzyme, telomerase, is responsible for the elongation of one strand of the telomeric DNA and compensation of its shortening throughout the replication process. Thus, telomerase activity functions as a propagation indicator; although such activity is not noticed in most somatic cells (exception of embryonic tissues, stem cells and reproductive organs). In about 80-90% tumour cells, telomerase activity is triggered and plays the role in telomere length elongation, which is used as a marker for the diagnosis of neoplastic transformation. Additionally, telomerase activity analysis is a useful tool in the search for telomerase inhibitors. All telomerase activity detection assays can be classified into three large groups based on direct detection of telomerase activity and expression, detection of telomere G-tail length, and detection of alternative lengthening of telomere.

1.10.1 Detection of telomerase activity and expression

Several methods have been devised to evaluate and measure the telomerase activity and expression of the telomerase components. The telomeric repeat amplification protocol (TRAP) is the most common method for detecting telomerase activity that allows one to perform semi-quantitative and quantitative analyses, using some modifications within the assay (*e.g.*, increasing rate of analysis, replacing the radioactive label by non-labelled compounds, and reducing the amount of side products, etc.) .[80, 246]

1.10.1.1 Telomeric Repeat Amplification Protocol (TRAP) Assay

The telomeric repeat amplification protocol can be split into three main steps, **a)** primer elongation, **b)** amplification of telomerase-synthesized DNA, and **c)** detection. At the elongation stage, telomeric repeats are added to the telomere-replicating oligonucleotide (TS, 5′-AATCCGTCGAGCAGAGTT-3′) by telomerase present in the cell extract. Then, PCR-amplification is carried out towards the telomerase-synthesized DNA with the use of specific primers (*i.e.*, telomere-replicating primers and reverse primers). At this stage, different labels (*e.g.*, radioactive, fluorescent labels) can be integrated into the telomerase-synthesized DNA. Afterwards, electrophoretic separation of PCR products and photographing (in the original TRAP assay) are carried out to detect and analyse the expression. Initially, in the TRAP assay, PCR products are identified in polyacrylamide gel through the radioactive label, using either a radioactively labelled primer or incorporated into the DNA during the PCR process. The method allows a qualitative assessment of the activity and enzymatic activity of telomerase in cell and tissue extracts. PCR at the second stage of the TRAP assay allows the acquisition of a sufficient amount of DNA for gel staining.

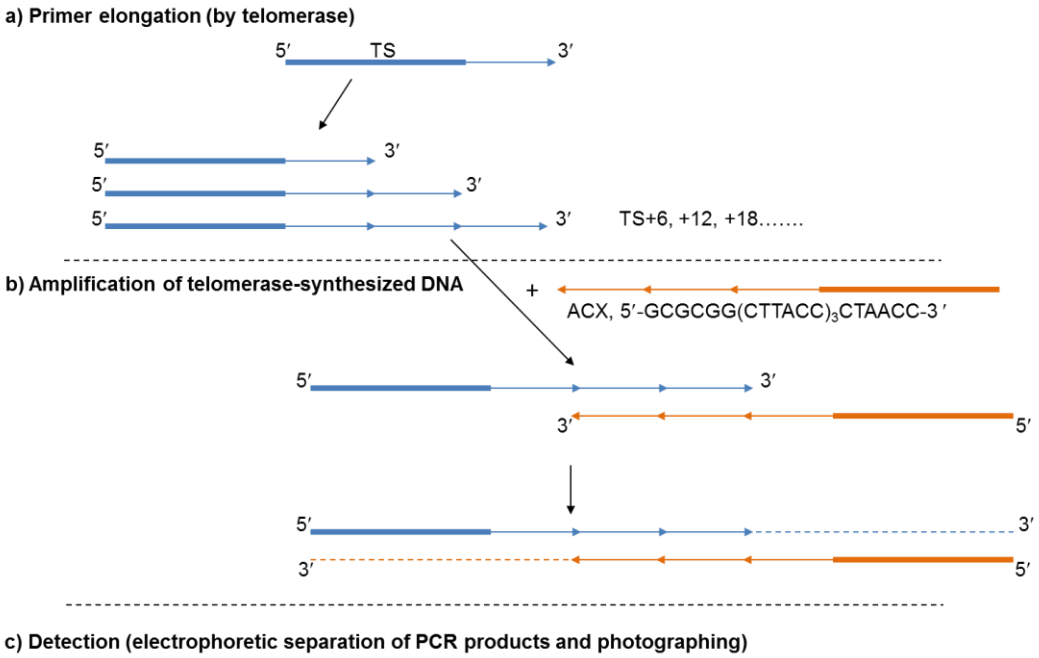


Figure 1.48: Summary of the original TRAP assay.

The original TRAP assay had some limitations. For example, use of the initial CX-oligonucleotide (*i.e.*, 5′-(CCCTTA)$_3$CCCTAA-3′, used in amplification of PCR products), which has a complementarily overlay with TS-oligonucleotide for several base pairs, results in the formation of dimeric primers and causes unaccepted products development. An ACX primer has been tried to overcome this problem, but a background signal appeared in a concentrated tumour extract after using even the optimal ACX primer with the non-complementary TS end.[247] Moreover, a problem occurred during the use of reverse primers (which is completely complementary to the telomeric repeats), where the primers were not annealed during PCR and caused a false signal. Generally, such a problem can be resolved by the addition of non-complementary nucleobases to telomeres with a 5′-end non-telomeric “adjunct” made of 6-bp to the primer ends. In order to decrease nonspecific signals, a combination of several reverse primers are used (*e.g.*, primers RP (5′-TAGAGCACAGCCTGTCCGTG-3′) and RPC3 (5’-TAGAGCACAGCCTGTCCGTG(CTAACC)$_3$-3′)), which is known as “two-primer” TRAP.[248] Gomez and co-workers reported that oligonucleotide TSG4 (*i.e.*, 5′-GGGATTGGGATTGGGATTGGGTT-3′) can also be added to the TRAP mixture to evaluate the effect of duplex-stabilizing inhibitors.[249] Moreover, during the usage of PCR in the signal

amplification process, the PCR inhibitors present within the specimen can affect the results in detecting telomerase activity.

Several advancements have been made in the TRAP assay, for example, a scintillation proximity assay (SAP) combined-TRAP assay where 5'-biotinylated oligonucleotides act as substrates and PCR amplification is carried out in the presence of tritiated (^{3}H) thymidine[250], a hybridization protection assay-TRAP where acridine bound probes are used to detect amplified DNA[251], a TRAP-ELISA method, where TRAP is combined with an enzyme-linked immunosorbent assay (ELISA)[252], a real-time quantitative TRAP assay where RT-PCR is used for simultaneous DNA amplification and measurement of the amount of products obtained after each amplification cycle[253], and an *in situ* TRAP assay where both substrate and reverse primers are labelled by FITC (*i.e.*, fluorescein isothiocyanate), then cell suspension samples are immobilized and dried, followed by incubation after addition of TRAP reagents and then *in situ* PCR is performed[254].

1.10.1.2 Direct detection of telomerase-synthesized DNA

Direct combination of a radioactively labelled substrate can be used to determine telomerase activity. In this method, the ability of telomerase can be electrophoretically assessed to elongate oligonucleotides in the presence of radio-labelled dNTP (*e.g.*, deoxy-guanosine triphosphate (dGTPl) as telomerase-synthesized telomeres are enriched in guanine bases). The main disadvantages of this method include the usage of large amounts of radioactive isotopes and in many cases observed inadequate sensitivity.[255]

Telomerase-synthesized DNA also can be identified in surface plasmon resonance (SPR). SPR demonstrates on a metal's surface under conditions of total internal reflection and is characterized by a generic angle of reflection, and consequently, the refraction index. As this effect shows itself on the metal surface, it spreads deep into the solution with exponential hindering as a distance function, where any interactions between the molecules can change the curbed wave and thus affect the characteristics of the surface plasmon, causeing a change in resonance angle and refraction index. Changes in refraction index are the characteristic measurements for evaluating the interaction between bio-molecules. With the help of a

biosensor, SPR is used to determine the activity of telomerase as the corresponding elongation of a telomere-replicating oligonucleotide.[256, 257]

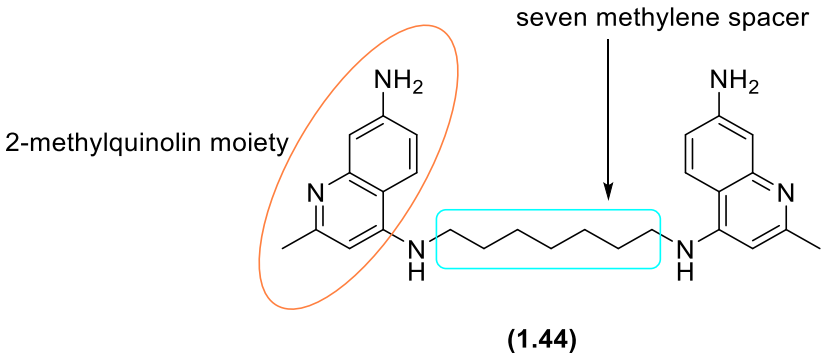
Moreover, telomerase activity can be determined using oligo-modified magnetic particles and NMR, where magnetic particles of an iron oxide are used to modify the oligonucleotides that are complementary to telomeric repeats.[258] According to the quartz crystal microbalance technique, an Au-quartz resonator is also used for a micro-gravitometric analysis of telomerase activity.[259] Ferrocenyl naphthalene diimide is used to detect the activity of telomerase electrochemically, which is based on the ability of telomeric G-quadruplex formation. In this method, ferrocenyl naphthalene diimide stoichiometrically binds with DNA quadruplexes and stabilizes the structure, and this bound ferrocenyl naphthalene diimide can be detected using electrochemical methods. Though this method cannot determine the length and position of individual DNA fragments, the sensitivity of this method is adequate to detect the telomerase-synthesized DNA without amplification.[260] FRET and total internal reflection fluorescence microscopy were employed to detect the telomerase activity by individual complexes of *Tetrahymena thermophila* telomerase.

1.11 Design and Rationale of this Project

1.11.1 Background

In 2012, Rahman and co-workers identified thirteen small molecules with varied chemical structures through screening 2307 molecules of NCI Diversity Set II, the NCI Natural Products Set II and the NCI Mechanistic Diversity Set libraries against human telomeric G-quadruplex (F21T, 5′d-GGG-TTA-GGGTTA-GGG-TTA-GGG-3′) using a FRET-based DNA melting assay.[261] These compounds (molecular weight range of 327-533 Daltons) showed significant selectivity towards the telomeric G-quadruplex compared with duplex DNA and the ΔT_m values were 7.3 to 24.2 °C at a 1 μM ligand concentration. Those molecules were also evaluated against a DNA/RNA hybrid duplex sequence (DRH, 5′-CCC-UAA-CCC-UAA -TTT-TTT-TTA-GGG-TTA-GGG-3′), where NSC 273829 (**1.44**) showed 11.5 °C and 3 °C stabilization at 5 μM and 1 μM ligand concentrations, respectively, by FRET melting assay (unpublished result). Moreover, molecule **1.44** was potent towards the F21T quadruplex sequence (11.8 °C at a 1

μM ligand concentration).[261] NSC 273829 is also a drug-like small molecule (log*P* = 4.60 and MW = 442.61 Daltons) and was also reported as active in mouse L1210 leukaemia xenografts[262] and against leishmania[263].



(1.44)

Figure 1.49: Structural features of NCI library compound NSC 273829.

1.11.2 Rationale of the Project

It has been reported in different articles that guanine-rich oligonucleotide sequences can fold into characteristic four-stranded G-quadruplex structures[145, 264] which are observed in telomeric DNA repeats and in promoter oncogene sequences and other regulatory regions of genes that are particularly involved in cellular proliferation.[153] However, the single-stranded 3′-overhang of the hTelo DNA sequence can be stabilized by small molecules to maintain G-quadruplex structures, thus preventing further annealing and subsequent addition of the telomerase enzyme complex (which is overexpressed in most cancer cells) to add further hTelo DNA repeats, and causing subsequent inhibition of cell growth and senescence.[265, 266] A number of hTelo sequences that form stable G-quadruplexes have been identified and characterised by NMR and crystallographic methods.[145, 147, 196, 267, 268]

Additionally, the single-stranded 3′-overhang of hTelo DNA forms a heteroduplex structure with the RNA-template of the telomerase enzyme complex during the catalytic cycle of telomerase expression in cancer cells, which is generally known as a DNA/RNA hybrid. It has already been established that the heteroduplex sequence forms during the initial annealing of the substrate telomere to the template RNA and during its extension in the 3′ direction.[107, 182] Such hybrid duplexes were discovered in 1960 and are appealing objectives for cancer

treatment due to their relatively unique structures in cancer cells compared with the DNA/DNA duplexes sequence[183], specificity in sequence (due to the telomeric repeat of TTAGGG) and precise length of the duplex (4-11 base-pairs)[269]. Most importantly, as the DNA/RNA hybrid duplex binds tightly to the protein portion of telomerase, molecules that bind to this hybrid duplex will also be able to access the unique protein surfaces of telomerase. Therefore, DNA/RNA hybrid duplex-binding molecules may be able to act as a platform for making precise analogues to identify the specific telomerase protein surfaces with improved affinity and specificity.[188] Moreover, the diversity in hybrid structures (due to the presence of different nucleobases (purine and/or pyrimidine) on the DNA strand) offers diverse backbone conformations and groove geometries in the DNA/RNA hybrid duplex that are distinctive features against which synthesised small molecule analogues can be selectively targeted.[184]

Small compounds that induce the formation of and/or selectively prevent dissociation of DNA/RNA hybrid duplex structures during the telomerase catalytic cycle or bind to G-quadruplexes are of interest to develop as potential therapeutic candidates, mostly in the anticancer drug discovery research area. The G-quadruplex binding agents that have been identified so far are mainly based on core functionalized polycyclic heteroaromatic systems with acridine, anthraquinone, perylene or porphyrin moieties.[201, 202, 270, 271] In general, relatively large aromatic ring systems are considered as a priority in G-quadruplex ligand design, since they can either interact with the terminal G-quartet plane through $\pi-\pi$ interactions (*i.e.*, end-stacking) or intercalate between G-quartets. On the other hand, there is no reported molecule so far that can target the telomeric DNA/RNA hybrid duplex structure either as a groove binder or intercalating agent. Thus this is a novel target in medicinal chemistry and cancer drug discovery research.

Several bulky aromatic functionalities have some shortcomings such as high binding affinity by themselves and reduced water solubility due to their hydrophobicity. To eliminate such problems, ligands were transformed into various analogues and functionalized by introducing one or more cationic side arms (for example, tertiary amine tail) to improve solubility and increase the chance of possible binding with phosphate backbones and/or with functional

groups available in the loops and grooves of DNA/RNA hybrid duplexes and/or telomeric G-quadruplex structures.

1.11.3 Aims and Objectives

The main aim of the project is to design and synthesize small-molecule ligands capable of interacting with telomeric DNA/RNA hybrid duplexes. Small molecules that bind to these heteroduplexes may inhibit telomerase activity by stabilizing the structure and thereby preventing strand dissociation or by distorting the substrate duplex leading to misalignment of the key catalytic groups. As telomerase is frequently overexpressed in cancer cells, the DNA/RNA hybrid duplex is considered a very specific and exciting therapeutic target. However, due to the unique structure of this heteroduplex, it is very challenging to target this sequence; very few studies have been carried out, and only a few compounds have shown DNA/RNA hybrid duplex stabilising properties.[100, 107, 184, 188, 191] In the present study, NSC 273829 (**1.44**) was selected as a lead compound based on our initial screening results and the rationale for this selection is discussed in section **1.11.1** and **1.11.2**. A series of analogues were designed based on initial modelling studies and these ligands have different structural diversities and conformational flexibilities.

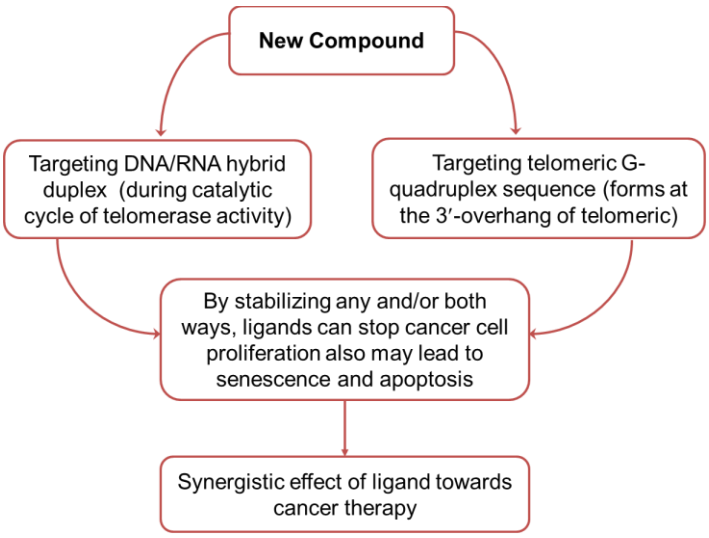


Figure 1.50: Expected outcome from this project targeting DNA/RNA hybrid duplexes to inhibit telomerase activity within cancer cells.

The specific objectives of the project are as follows:

1. Synthesise novel DNA/RNA hybrid duplex and/or G-quadruplex ligands using a variety of chemical approaches.
2. Evaluate new molecules using FRET-based DNA thermal denaturation assays to evaluate the affinity and selectivity of newly-synthesized ligands towards particular DNA sequences (DNA/RNA hybrid duplex and hTelo sequence, F21T). For all the experiments, duplex DNA sequences were used as controls for the telomeric DNA sequences (F21T), whereas cDRH (where RNA bases of DRH were replaced by corresponding DNA bases) was used as a control for DNA/RNA hybrid duplex sequences.
3. Evaluate the novel ligands for cytotoxicity in tumour cell lines.
4. Use the results to compare the newly-synthesized ligands with those previously synthesized to derive structure activity relationships (SARs). SARs should contribute to a greater understanding of the interactions of newly-synthesized small molecules with G-quadruplex structures and/or DNA/RNA hybrid duplexes.

Chapter 2: Methods and Materials

2.1 Sourcing of Chemicals

In this research project, all chemical reagents and solvents were purchased commercially from Sigma-Aldrich, Maybridge, Acros, Alfa Aesar, TCI Europe, Fluorochem, Biofarma, Fisher Scientific and Fluka.

2.2 Analytical Tools

The liquid chromatography–mass spectrometry (LC-MS) technique was applied to monitor reaction progression and identification of the compounds. LC-MS was performed on a Waters Alliance 2695 with water and acetonitrile as the mobile phases. Formic acid (0.1%) was used with the acetonitrile to confirm acidic conditions during the course of analysis. The gradient conditions were acetonitrile/water (95%) for 2 minutes, which was then increased to 50% acetonitrile over 3 minutes. The gradient was then held at 50% acetonitrile for one minute and then increased to 95% acetonitrile over 1.5 minutes. The quantity of acetonitrile was then returned to 5% over 1.5 minutes and held for 0.5 minutes. The total duration of each run was 10 minutes. The flow rate was 1.5 mL/min; 200 µL was split *via* a zero dead volume T-piece which passed into the mass spectrometer. The wavelength range of the UV detector was 220-400 nm. A diode array (535 scans) was functionalised with the system. A monolithic (C-18, 50X4.60 mm) column was used.

Proton (^{1}H) and carbon NMR (^{13}C) spectroscopy were carried out on a Bruker Avance 400 MHz spectrophotometer. Chemical shifts (δ H) are cited in ppm (parts per million) and referenced to deuterated chloroform ($CDCl_3$, residual signal ^{1}H δ = 7.26, ^{13}C δ = 77.2), deuterated dimethyl sulfoxide (DMSO-d_6, residual signal ^{1}H δ = 2.54, ^{13}C δ = 40.45), deuterated methanol (MeOD, residual methanol signal ^{1}H δ = 3.31, ^{13}C δ = 49.00). Multiplicities in ^{1}H NMR spectra are quoted as s = singlet, d = doublet, t = triplet q = quartet, m = multiplet, dd = doublet of doublets, ddd = doublet of doublet of doublets, dt = doublet of triplets, td = triplet of doublets, spt = septet and br = broad. The code (0) in ^{13}C NMR spectra denotes the presence of a quaternary carbon.

High resolution mass spectra (HRMS) were obtained on a Thermo Navigator mass spectrometer coupled with liquid chromatography (LC) using electrospray ionisation (ES) and time-of-flight (TOF) mass spectrometry. Infrared spectra (IR) were recorded on a Perkin Elmer spectrum 1000 instrument.

Every single reaction was checked by analytical thin-layer chromatography (TLC) performed on E. Merck silica gel-60 F_{254} layered plates (0.25 mm). TLC plates were visualized under UV light (254 or 360 nm) and/or by staining the plates with either vanillin spray or potassium permanganate solution followed by heating.

2.3 Purification Techniques

2.3.1 Flash Column Chromatographic Technique

Compounds were purified by traditional column chromatography using Merck Flash Silica Gel-60 (230-400 mesh). To get an idea of a proper solvent system, traditional TLC was utilised using different solvent systems to obtaine the best separation profile.

Automated flash column chromatography is an air pressure-driven hybrid of medium pressure and short column chromatography, optimized for rapid separations on the basis of UV and ELSD detection. A Reveleris® Purification System was used in this project to purify three compounds, where the mono-adduct and bi-adduct products' retention times (*rf* value) were very close to each other.

2.3.2 Catch and Release Method

SCX-2 resin cartridges were purchased from Biotage Synthesis (Virginia, USA) with different sorbent capacities (500 mg, 1 g, 2 g, 5 g and 10 g). The cartridges contain propylsulfonic acid-functionalized silica. The strong cation exchange sorbent is a very good choice for basic drug extraction. The cartridges were first activated by an initial wash with 2 column volumes of dichloromethane and 4 column volumes of methanol. The reaction mixtures were then poured onto the cartridge (sorbent mass should be 10 times that of the calculated mass of the product) and the solvent was allowed to pass through the cartridges under gravity. The cartridges were then washed with dichloromethane (3 times), dimethylformamide (3 times) and methanol (1

time) and this cycle was repeated three times to remove the impurities, using a vacuum to accelerate the elution. Products were released from the cartridge using 2M ammonia solution in methanol and concentrated *in vacuo*.

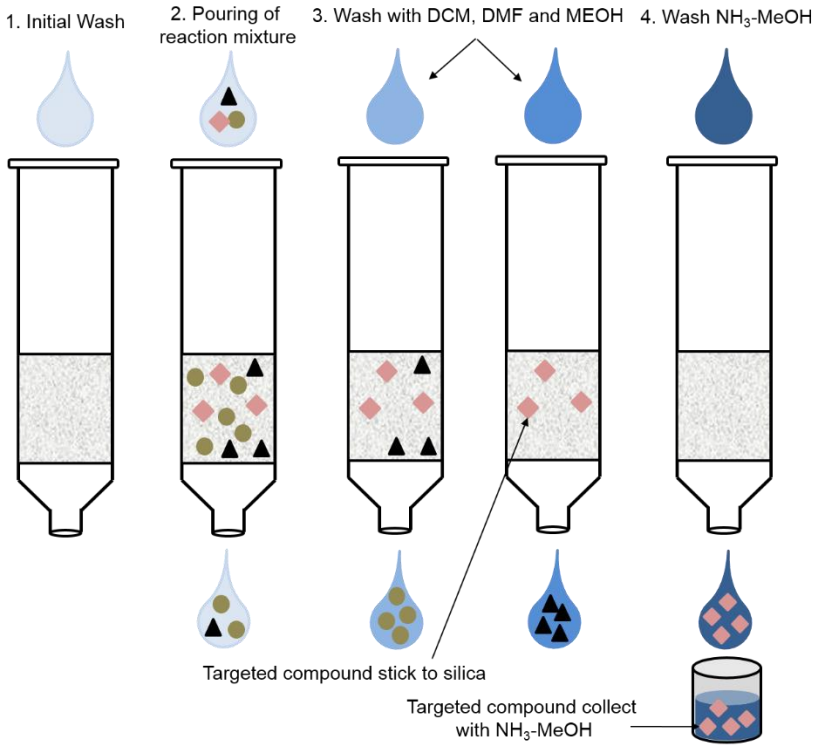


Figure 2.1: Procedure for the 'Catch and Release' purification method.

2.3.3 Trituration Technique

This process was used to purify crude chemical combinations with solid impurities. A solvent (either polar or non-polar) was selected in which the desired product was very soluble and the unwanted by-products were insoluble (or *vice versa*). The crude material was washed with the solvent and filtered away, leaving the purified product in solid form and any impurities in solution. In this project, a few compounds were purified through triturating the crude sample with diethyl ether (20-30 mL) to give a white solid which was then filtered and washed with diethyl ether to give the pure product.

2.4 General Amide Coupling Reaction

In this project, a standard method for amide coupling was employed to prepare all *bis*-amide molecules in libraries 1 to 7. The reaction was carried out using 1-hydroxybenzotriazole (HOBt)

and diisopropylcarbodiimide (DIC) and the mechanism for this reaction is illustrated in **Figure 2.2**.[272] The HOBt/DIC-mediated amide coupling reaction was employed for all reactions due to the availability of the reagents and previous use within the research group with good yield. Additionally, most of the reactions were made with either dimethylformamide (DMF) or dichloromethane (DCM) as the solvent, because the urea by-product is soluble in DMF (easy to remove during washing) and the initial activation step is faster in nonpolar solvent like DCM.[273, 274]

Initially, the activated ester was generated by treating the corresponding acid with HOBt in the presence of DIC.[275, 276] After its formation, selected diamines were added and in most cases, the reaction was completed after 14-18 h.

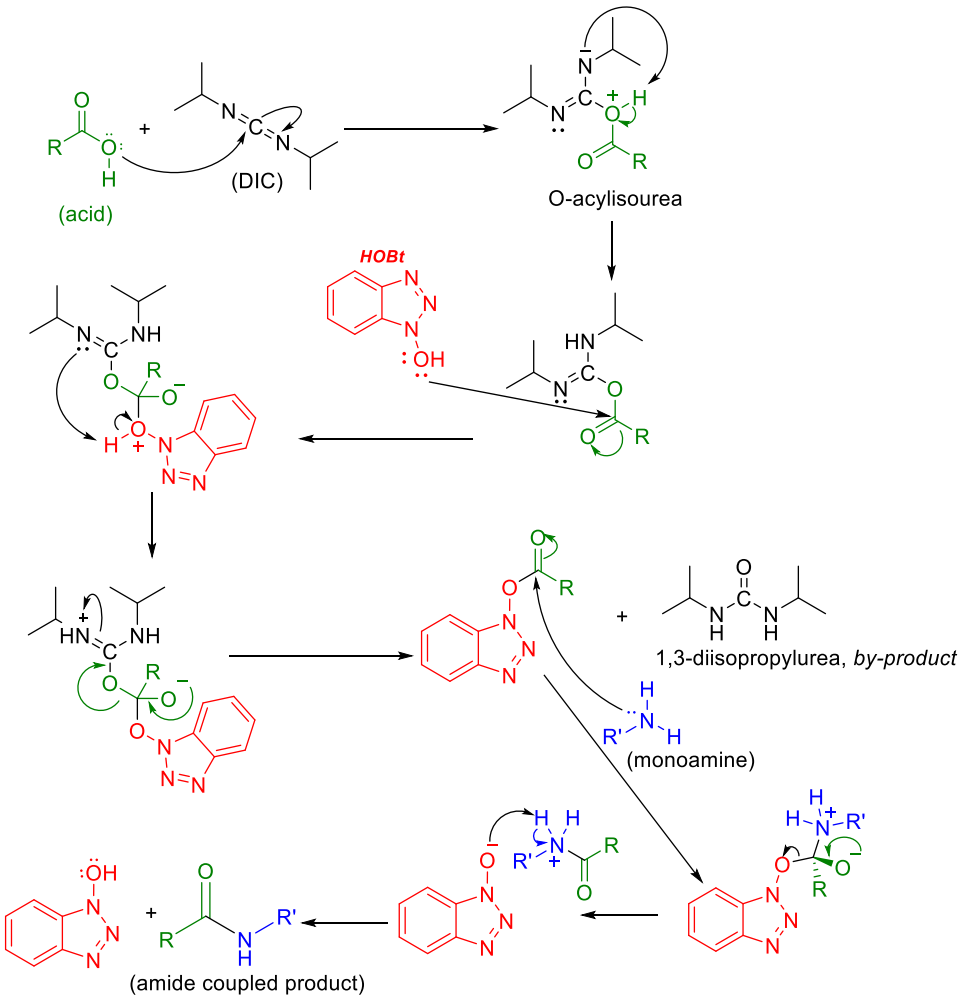


Figure 2.2: Schematic diagram of the HOBt/DIC-mediated amide coupling reaction.

2.5 Methods Used for Biophysical and Biological Evaluation

2.5.1 Fluorescence Resonance Energy Transfer (FRET) Assay

In 1970, the Fluorescence Resonance Energy Transfer assay, a quantum phenomenon, was first applied in biological research, especially in proteins. The FRET assay was first demonstrated by Theodor Förster over 70 years ago.[277] The principle of this assay is the excitation of a fluorophore and transfer of energy from the donor to an acceptor fluorophore over dipole-dipole interaction without any discharge of photons. Consequently, the donor molecule's fluorescence is quenched, and the acceptor molecule becomes excited, and finally it loses energy *via* heat or fluorescence emission (also known as sensitized emission) (**Figure 2.3**). Such resonance transfer occurs over more than interatomic distances, without conversion to thermal energy and any molecular collision. This transfer of energy leads to a fall in the donor's fluorescence intensity and excited state lifetime, and conversely causes an increase in the acceptor's emission intensity. A pair of fluorophores that interact in such a way that transfer of fluorescence resonance energy occurs is referred to as a donor/acceptor pair.[278]

Remedios and co-workers reported that 'Fluorescence Resonance Energy Transfer' is a spectroscopic process by which energy passes in a non-radiative way between fluorophores over distances of 10-100 Å.[279]

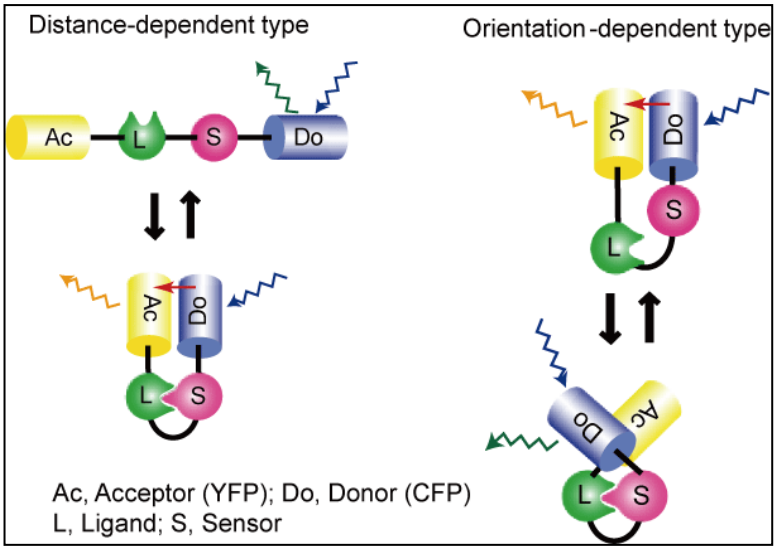


Figure 2.3: Principle of the FRET mechanism; the 'donor' fluorophore (Do) absorbs a photon and transfers this energy by a non-radiative process to the 'acceptor' fluorophore (Ac).[280]

In theory, the donor fluorophore can directly transfer its excitation energy to the acceptor fluorophore through distant dipole-dipole intermolecular coupling in the case of overlaying of the fluorescence emission spectrum of the donor fluorophore onto the absorption spectrum of the acceptor fluorophore, and if the two fluorophores are within a minimal spatial radius. The acceptor molecule quenches the donor molecule's fluorescence during the energy transfer, and if the acceptor itself is a fluorophore then amplified or sensitized fluorescence emission is observed. The phenomenon is detected by exciting a specimen containing both donor and acceptor fluorophores with light of wavelengths corresponding to the absorption maximum of the donor fluorophore, and detecting light emitted at wavelengths centred near the emission maximum of the acceptor.

For effective FRET to occur, the donor and acceptor fluorophores need to be in close proximity to each other and the absorption or excitation spectra of the acceptor fluorophore must overlay the fluorescence emission spectrum of the donor fluorophore. Förster established that the efficiency of the process (E) depends on the inverse sixth-distance between donor and acceptor as follows:[277]

$$E = \frac{R_0^6}{R_0^6 + r^6}$$

Here, R_0 = Förster distance, r = Actual distance between donor and acceptor fluorophores

It is already established that stabilization of the DNA quadruplex and/or duplex helical structure due to the addition of a small drug-like molecule results in elevated melting temperature of the ligand-bound oligonucleotide compared with the oligonucleotide itself. As a result, the difference in melting temperature (ΔT_M) of the ligand-bound oligonucleotide and oligonucleotide itself gives an indication of the stabilizing efficiency of the molecule to the DNA structure. Choice of buffer is an important factor to obtain accurate melting temperatures. Usually most real time PCR works within the range of 25-100 °C, and, therefore, the initial melting temperature of the quadruplex and/or other sequences used should be within the range of 45-55 °C that can be adjusted by using an appropriate buffer.[281] Usually, a lower melting temperature (T_M) will lead to imperfect folding of DNA at low temperatures, whereas a higher T_M will prevent the analysis of extremely potent ligands. Concerning the effect of pH, it is important to note that the properties of fluorescein (FAM)-conjugated oligonucleotides are pH-dependent as pHs lower than 7.0 can cause an important quenching.[282, 283]

Generally, the donor and acceptor fluorophores are different and FRET can be detected either by the appearance of acceptor fluorescence or by quenching the fluorescence of the donor fluorophore. A fluorescent molecule is always chosen as the donor moiety. With appropriate energy of excitation, electrons from the donor fluorophore jump from the ground state (S_o) to an excitation level and these electrons come back to the lowest vibrational state (S_1) within a very small time space (can be picoseconds) and ultimately fall back to the ground state (S_o) by releasing a photon of light. At the point of fluorescence resonance energy transfer, there is competition between the decay of the donor fluorophore and energy transfer to the acceptor fluorophore for the falloff of excitation energy. With resonance energy transfer, the energy is transferred to the acceptor fluorophore rather than the transfer of a photon and those electrons become excited as stated for the donor fluorophore, consecutively. Finally, the successive return to the ground state releases photons (**Figure 2.4**).

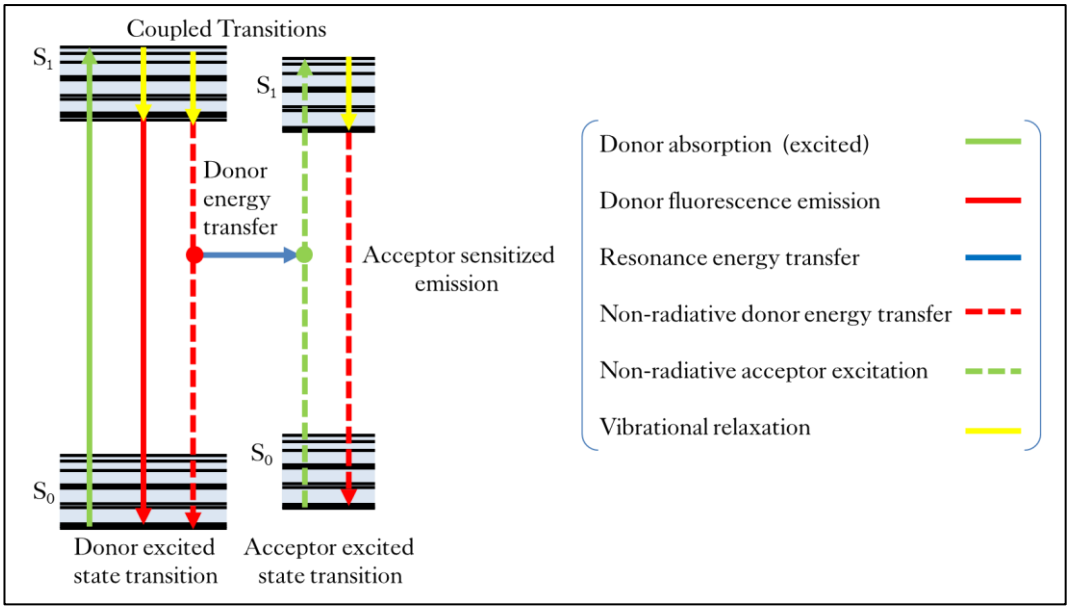


Figure 2.4: Jablonski diagram illustrating the coupled transitions involved between donor emission and acceptor absorbance in FRET.

A 'FRET melting assay' utilizes this basic principle where an increase in temperature leads to denaturing or melting of the oligonucleotides, causing an increase in distance (Förster distance) between the two fluorophores and ultimately leading to a rise in fluorescence energy. This technique has been found to be very useful in investigating various types of nucleic acid structures.[284] However, availability of suitable labelled substrates limits its use in biological systems. However, this is not a problem for studying quadruplex DNA stability as synthetic DNA is usually used for all analysis relating to G-quadruplex formation and labelled oligonucleotides can be obtained easily.[285]

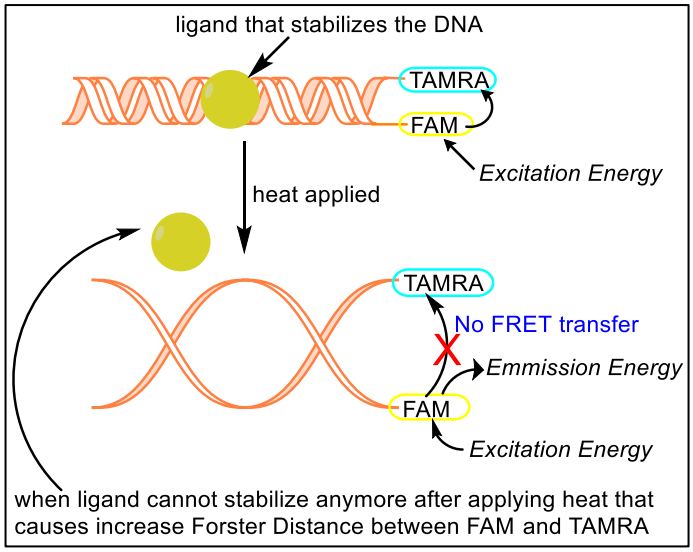


Figure 2.5: FRET transfer mechanism and application to a ligand on a DNA duplex sequence.

In 2001, Mergny and co-workers first used the FRET technique to investigate the G-quartet structures.[286] A synthetic oligonucleotide containing a pair of fluorophores on both the 5' and 3' ends with a suitable Förster diameter that followed the principle of the fluorescence resonance energy transfer mechanism was used so that, upon folding, the donor and acceptor fluorophores were brought close to each other to allow efficient fluorescence resonance energy transfer.

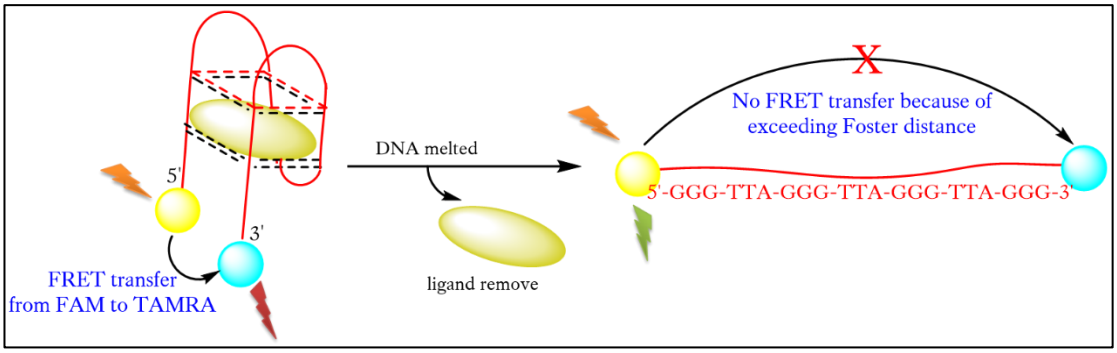
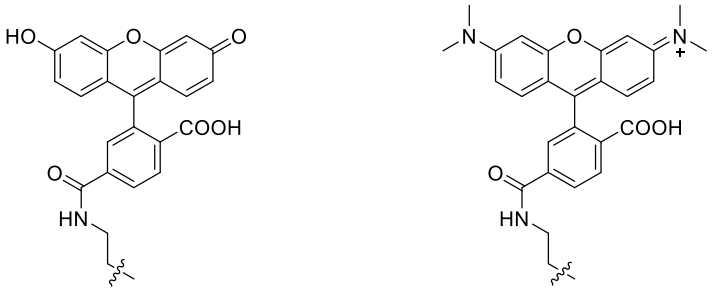


Figure 2.6: Structure of the dual probe-labelled quadruplex, stabilized by a ligand (oval shaped) and subsequent schematic representation of FRET-based experiments.

FRET gives valuable evidence about nucleic acid structures[287], as the mechanism is distance- and orientation-dependent[277]. Considering multi-stranded nucleic acids structures, the fluorescence resonance energy transfer method has been effectively applied in the case of triple helices[288], C-rich quadruplexes (i-motif)[283] and c-Myc quadruplex sequences.[282]

To conduct the FRET melting assay in our study, 6-carboxyfluorescin (FAM) and 6-carboxytetramethylrhodamine (TAMRA) were used at the 5' end of oligonucleotide sequences as a donor and at the 3' end of oligonucleotide sequences as an acceptor dye, respectively.



(**2.1**) (**2.2**)

Figure 2.7: Structure of FAM (**2.1**) and TAMRA (**2.2**)

In principle, at low temperatures the oligonucleotide remains folded to keep the probes within the Förster radius (R_0) resulting in quenching of the emission radiation of FAM by TAMRA (*i.e.*, which were used in this project as a FRET pair). The peak absorbance wavelength for FAM is 494 nm with a peak emission wavelength at 518 nm. If FAM and TAMRA are tied at the 5'- and 3'-ends respectively of an oligonucleotide and this construct is excited at 494 nm, as long as the oligonucleotide remains intact, emission will be at 580 nm and not at 518 nm as FAM is transferring its energy to TAMRA. As soon as the oligonucleotide is disordered (*i.e.,* disrupted) by either heat or an exonucleolytic reaction excitation at 494 nm will result in emission at 518 nm. This is due to the fact that the pair is no longer bound and falls apart from Förster distance (R_o). However, an increase in temperature toward the melting point of the oligonucleotide leads to unfolding of the oligonucleotide and causes the fluorophores to be separated sufficiently that the FRET efficiency is ultimately diminished. The FRET technique is a simple, quick, convenient, and cost effective method that can be conducted in a variety of ionic conditions and can easily be adjusted for 96-well plate high-throughput screening. For

these reasons, it has been adopted by various research groups and a large number of molecules with diverse structural features have been screened using this technique. However, this method has some disadvantages as well; sometimes it can give false positives due to quenching of the donor emission by chromophores present in the ligands and sometimes it can produce false negatives because of folding of the DNA sequence with the ligands in such a way that the distance between the donor and acceptor fluorophore prevent the effective transfer of fluorescence resonance energy. To avoid the problem of false positives, the absorbance for the ligands should be checked beforehand to make sure that they do not absorb at the same wavelength as the acceptor. The DNA bases, especially the guanines, can also cause quenching of the donor emission, hence they should be separated from the donor moiety.

2.5.1.1 Procedure involved in FRET Assay

2.5.1.1.1 Selection of oligonucleotides

The synthesised molecules were screened against the following oligonucleotides in this study:

DRH; Telomeric DNA/RNA hybrid duplex – A hairpin DNA/RNA hybrid duplex was used in this experiment where the DNA strand corresponds to the human telomeric repeat sequences and the RNA strand corresponds to the RNA template of telomerase (**Figure 2.8**).

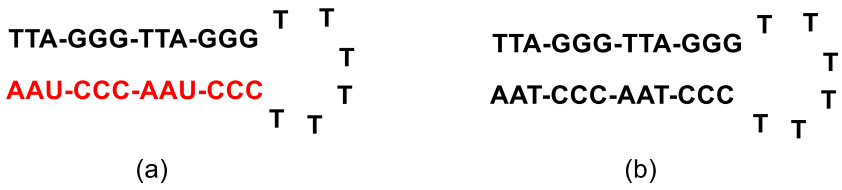


Figure 2.8: (a) Hairpin structure of telomeric DNA/RNA hybrid duplex (**red** indicates the RNA nucleotides) and (b) Hairpin structure of control cDRH duplex.

cDRH; Control DNA/DNA duplex – A hairpin form of oligonucleotide was used as a control, which has an identical sequence and orientation, except that the RNA nucleotides are replaced with the corresponding DNA nucleotides (**Figure 5B.1 b**).

F21T; telomeric quadruplex DNA – For quadruplex ligands to function as telomerase inhibitors, they must show significant quadruplex-stabilizing ability. A double dye oligonucleotide sequence with human telomeric repeats (F21T; 5′-FAM-GGG-TTA-GGG-TTA-GGG-TTA-GGG-TAMRA-3′) was used in this experiment.

cDDS; control duplex DNA sequence – A randomly chosen double dye-containing DNA was used as a control for the whole experiment. The sequence that was used was comprised of 45 DNA bases (5′-FAM-TAT AGC TATA TTT TTT TATA GCT ATA-TAMRA-3′).

All oligonucleotide sequences were purchased from Eurogentec, UK with a confirmation of RP-HPLC purification and label of product consistency.

2.5.1.1.2 Methods Adopted for FRET Melting Assay

Preparation of FRET-Buffer: 500 mL of buffer solution was prepared with potassium hydroxide and potassium chloride and adjusted to a pH of 7.4 with cacodylic acid (50 mM and 500 mM). The buffer was stored in the freezer at -20 °C.

Annealing of DNA: The fluorescence-tagged oligonucleotide sequences were diluted with sterile DEPC water (DNA Grade, Fisher Scientific) to obtain 20 µM solutions. 400 nM solutions prepared by serial dilution using FRET-buffer were annealed by heating at 85 °C for 5 minutes followed by cooling to room temperature over 3-4 hours (Grant Bio PCH-2 Dry Block Heating/Cooling System).

Preparation of Ligand Solutions: For all ligands, 5 mM stock solutions were made with dimethyl sulfoxide (≥ 99.9 %, A.C.S. spectrophotometric grade, Sigma-Aldrich). From these stock solutions, 100 µM, 10 µM, 4 µM and 2 µM working solutions were prepared by serial dilutions with FRET buffer.

Preparation of Plates: For the FRET-based DNA melting assay, 50 µL of annealed DNA was added to each well of a 96-well plate (Bio-Rad Laboratories), then 50 µL of ligand solution was added . Three different concentrations (10 µM, 4 µM and 2 µM) of each compound were tested in triplicate. Most commonly, pure FRET buffer instead of ligand solution was added to all wells of the first line (A) of the plate to serve as a blank. After 15 minutes of incubation at room

temperature, the plate was processed in a DNA Engine Opticon (Continuous Fluorescence Detector, MJ Research) and florescence measurements taken over a temperature range from 30-100 °C at intervals of 0.5 °C. Prior to each measurement, the temperature was kept constant for 30 seconds. The incident radiation was emitted at 450-495 nm and detection measured at 515-545 nm.

Data Processing: The achieved data were analysed using the program Origin – Scientific Graphing and Analysis Software (Version 7.0, OriginLab Corp.). The increases in melting temperatures (ΔT_m) were obtained by subtracting the value of the blank from the measured values of each sample. For each concentration of every compound, the average ΔT_m was calculated from the three corresponding values and plotted against the concentration of the ligand.

2.5.1.2 Melting curve analysis

Melting curve analysis examines the dissociation of dsDNA upon heating. As the temperature is increased, the double strand dissociates, leading to an increase in absorbance (*i.e.*, hyperchromicity). The temperature at which 50% of the DNA is denatured is known as the melting point. This melting can be detected using the FRET assay, in which fluorophores are attached facing each other at the 3′ and 5′ -end of the complementary strands (for an intermolecular duplex) or at the opposite ends of an intramolecular (hairpin) duplex. As the temperature is increased there is an increase in the absorbance (or fluorescence), generating a melting curve.

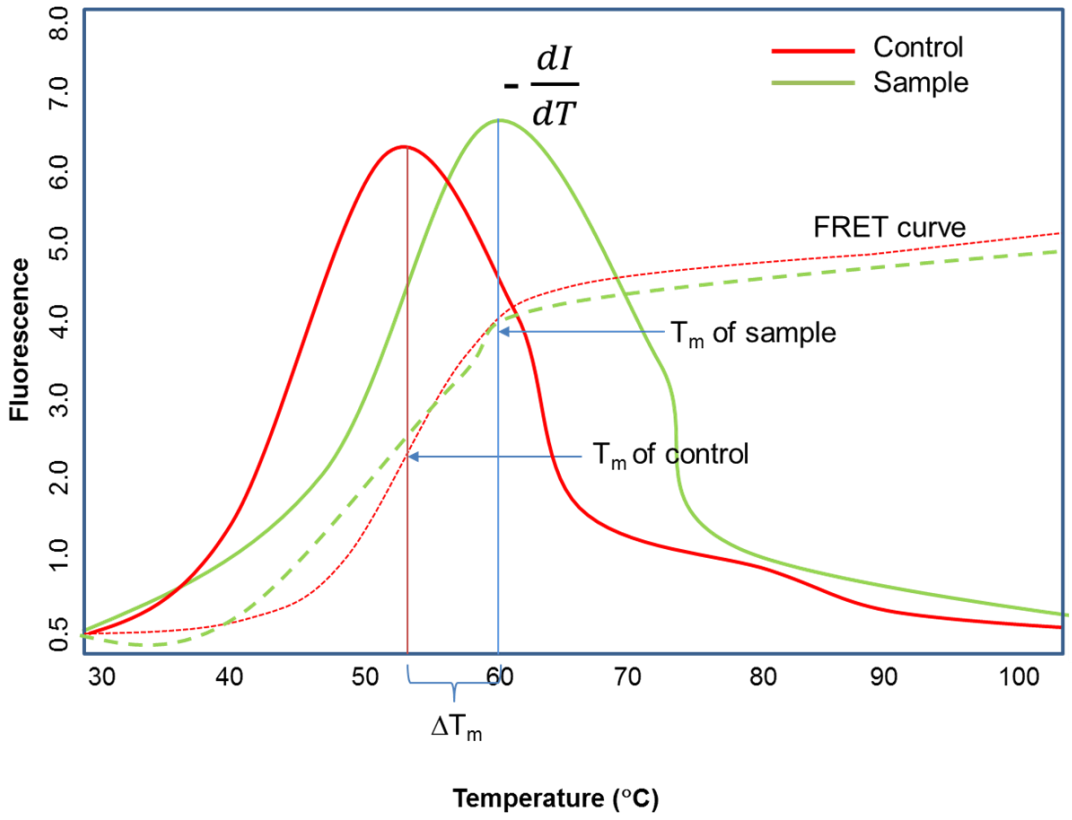


Figure 2.9: Typical dissociation curve of DNA melting.

2.5.1.3 Limitation of FRET melting assay

Sometimes the FRET melting assay can give false positive results due to quenching of the donor emission by the planar chromophore present in the ligands, and sometimes can produce false negative results because of folding of the oligonucleotide with the ligands in such way that the distance between the donor and acceptor prevent efficient FRET from occurring. To avoid the problem of false positives, the absorbance of the ligands should be checked prior to the assay to make sure that they do not absorb at the same wavelength as the acceptor. The DNA bases, especially the guanine, can also cause quenching of the donor emission, and hence it should be separated from the donor moiety.

2.5.2 Molecular Modelling and Molecular Dynamics (MD) studies[1]

Molecular modelling comprises theoretical methods and computational techniques together that can be employed to model or imitate the behaviour of molecules towards specific DNA sequences or protein structures. The system is used in the areas of computational chemistry, new drug design and computational biology to study molecular systems within the body, where molecules from small chemical scaffolds to bulky biological molecules can be analysed. The general feature of molecular modelling and molecular dynamics techniques is the explanation of a chemical scaffold at an atomistic level.

Molecular modelling techniques allow for the atomic-level evaluation of shape-fit and binding affinity of ligands with a receptor, and can be useful tools with which to establish structure-activity relationships of libraries of ligands and to drive the drug discovery process. Molecular docking calculations involve the use of a static crystal structure, and are particularly useful in determining shape-fit in a receptor. Molecular dynamics simulations, on the other hand, consist of the calculation of the time-dependent behaviour of molecular systems, and have provided valuable information on the changes in conformations of biomolecules (for example proteins or nucleic acids) as predicted over a certain time course.

The fields of *in silico* modelling and molecular dynamics have progressed rapidly through technological advancement. The speed of computers has doubled every two years ('Moore's Law' [289]) for the past number of years, and the use of specialized GPUs and supercomputers has allowed the docking of large libraries of molecules and simulation of large macromolecules. This has culminated in the 1ms all-atom molecular dynamics simulations of BPTI[290] using a specially built, massively parallel supercomputer capable of simulations of orders of magnitude greater than standard computer systems[291].

Statistical mechanics is central to the concept of molecular dynamics simulations and it is the study of microscopic systems to produce macroscopic conclusions. In MD simulations, individual atoms within systems are simulated in order to determine features of the system. Ensembles are defined as systems which have different microscopic states, but identical

[1] Molecular modelling and molecular dynamics studies were conducted by Dr Paul J M Jackson and Mr Meir Touitou of IPS, King's College London.

macroscopic ones, and in order to simulate whole systems, averages corresponding to experimental observables are defined in terms of 'ensemble averages'. The central hypothesis to this is that over the course of a simulation, all possible macroscopic states should be simulated by different microscopic states, and, therefore, at the end of a simulation, the system will have passed through all states. From a drug discovery perspective, the simulation of a ligand with a receptor should result in the determination of the lowest energy conformation (and thus, most biologically likely) combined drug-receptor complex structure.

In the case of DNA-RNA hybrid drug discovery, a molecular dynamics-driven approach was taken, as this allowed for the evaluation of different mechanisms of binding of the ligand library (*e.g.*, DRH minor groove binding or intercalation). As the main mechanism of binding of G-quadruplex-binding agents is through shape-fit and stacking interactions at the G-tetrad interface, a molecular docking-based approach was taken to assess ligand binding and correlate with FRET melting temperature experiments.

2.5.2.1 Process involved in molecular modelling experimentation for DNA/RNA hybrid duplex

The DNA/RNA hybrid duplex was constructed by first using *make-na* (http://structure.usc.edu/make-na/server.html) to create the RNA strand, followed by construction of the DNA sequence (including the TTT loop) using the AMBER module *nab*. The TTT loop was then covalently linked to the DNA backbone using parameters derived in-house. All compounds were docked in the minor groove of the DNA/RNA hybrid duplex sequence using AMBER *xleap*, parm99SB and modified parmbsc0 and Gaff AMBER force field parameters. Energy minimization was then undertaken in a gradient manner by initially placing the DNA or DNA/RNA hybrid duplex under a high force constraint (*i.e.* 500 kcal mol^{-1} $Angstrom^{-2}$), which was reduced in stages to zero to enable the ligand to find its local energy minimum, followed by reduction in force in a periodic manner with a relaxation of restraints. Once in equilibrium, production simulations were run for a period of 10 ns, and atomic coordinates were saved at 1 ps intervals. Simulations involving the control DNA duplex were undertaken in an identical manner with the use of the AMBER module *nuc* to create the DNA

sequence. Analysis of molecular dynamics simulations was undertaken using *VMD*[292], and all models were created using *Chimera.*[293]

2.5.2.2 Process involved in molecular modelling experimentation for hTelo Quadruplex sequence

Receptor Preparation: The F21T quadruplex structure was downloaded from the Protein Data Bank (PDB ID: 3CDM). The quadruplex was then subjected to a number of preparation steps in *Chimera,* including assignment of partial charges (AMBER ff98SB) and removal of hydrogens using the *DockPrep* module of AMBER, followed by writing to mol2 and pdb files. A molecular surface of the receptor was generated using *write dms*.

Ligand Preparation: All ligands were constructed and energy minimised using ChemBioOffice and exported in pdb format and converted into the Sybyl format mol2 using *Chimera*.[293] The *DockPrep* module in *Chimera* was then used to assign partial charges to each molecule (*AM1-BCC* charges in this instance), and atom types were subsequently assigned *via* the AMBER GAFF force-field using ANTECHAMBER.

Docking: Docking experiments were undertaken using the DOCK6 software suite and the F21T quadruplex (PDB ID: 3CDM) was used in the study. In each case, the quadruplex interface was used as the receptor and the ligand used in the crystallography study was used as a guide for the docking site.

The receptor was prepared (outlined above), and a number of steps were undertaken in DOCK6 to isolate the binding pocket of interest. Firstly, spheres were generated around the surface of the molecule using *sphgen* and *Sphere_selector* was then used to filter results. Spheres within 15Å of the ligand were selected for further analysis. This resulted in the assessment of the full quadruplex.

Finally, every ligand was automatically positioned into the spheres with the maximum number of conformations set at a high level (500) to explore a large amount of conformational space, thus producing a docked ligand/DNA structure.

Evaluation of Ligand Binding: The ligand was evaluated based on the DOCK scoring function MMGBSA. During the docking process, a grid was created around the receptor. The grid was then used to allow rapid score evaluation in DOCK. Prior to scoring, orientations of the ligand which exhibited significant steric interactions with the receptor molecule were discarded using the *bump* filter.

In evaluating ligand/DNA interactions, two factors were considered crucial; shape-fit of the molecule to the DNA receptor, and positioning of the molecule on the DNA-binding interface. The latter was considered particularly relevant as in order to stabilise the quadruplex structure, it is known that ligands must bind to the G-tetrad.

2.5.2.3 Limitations of MD simulation studies

A number of limitations are associated with the molecular modelling techniques used in this study. In the case of the docking studies, the crystal structure used in the methodology (PDB ID 3CDM) represents a structure of the F21T quadruplex. This ligand is not related in any way to the ligands synthesised within this study, and until a ligand:DNA complex is crystallised containing one of the analogues synthesised within this study, the docking results can only provide an estimate of location of binding. Similarly, a limited number of conformations of the ligand were explored during docking, and only a single conformation of the quadruplex DNA is explored in this type of study. Molecular dynamics simulations, on the other hand, explore multiple conformations of both the ligand and DNA and provide a far greater representation of the ligand:DNA complex. However, in a similar manner to docking studies, the positioning of the ligand on DNA can only be regarded as an estimate.

In the case of DRH, every effort was made to generate a conformation of the molecule which was biologically relevant. Although a crystal structure of the receptor was not available, an A-form structure containing both DNA and RNA was generated, and then energy minimised. This provided a close approximation to the proposed biological structure of the DRH receptor. Conformation analysis of the ligand suggested that the lowest energy conformation of the structure was a chair-like conformation, and this conformation was then used for MD studies. As is the case with docking studies, this, too, is an estimate of the location of binding, and if a

crystal structure were available, this would provide confirmation of our estimates generated during MD simulations.

2.5.3 Circular Dichroism (CD) Spectroscopy

Circular Dichroism (CD) is detected while optically active molecules absorb left and right hand circular polarized light in a different way to some extent. Dichroism is a property of absorbing light to different magnitudes depending upon the polarization form of the incident stream of light; if the absorption of light circularly polarized in a right-handed direction is different from the absorption of light circularly polarized in the opposite direction (left handed), then the molecule will exhibit circular dichroism.[294] CD spectra for different types of secondary structure existing in peptides, proteins and/or nucleic acids are not the same due to a varied Cotton effect (the typical variation in optical rotatory dispersion and/or circular dichroism in the surrounding area of an absorption band for a molecule) of the macromolecules. Therefore, the analysis of CD spectra can provide useful evidence about the secondary structures of biological macromolecules and their conformations.[295]

However, we are not aware of any reports of CD analysis of telomeric DNA/RNA hybrid duplexes to date, which is a crucial hybrid sequence that forms during the catalytic pathways of telomerase expression. In 1999, Ren and Chaires screened some standard groove binding agents and intercalators against thirteen oligonucleotide sequences, where they established a CD spectrum of a RNA/DNA hybrid sequence (*i.e.*, poly(rA)-poly(dT)).[185] According to the spectrum (**Figure 2.10**), a positive CD signal was observed at around 260 nm and a sharp negative CD signal at around 248 nm.

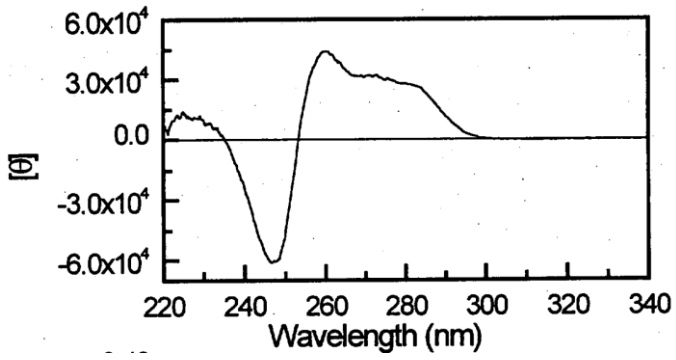


Figure 2.10: CD spectrum of RNA/DNA hybrid (*i.e.*, poly(rA)-poly(dT)) in BPES buffer.

In 2009, Kypr and co-workers reported that the CD spectrum of A-form DNA of heterogeneous structure is generally the same as that of the A-form of the corresponding RNA or DNA-RNA hybrid with the same primary structure, which is characterized by a dominant positive signal at 260 nm and a negative signal at 210 nm (where the amplitude of the positive signal is typically within 7-12 $M^{-1}cm^{-1}$ depending on the base sequence).[296] In 2012, Xu and colleagues suggested that the course of DNA/RNA G-quadruplex structure formation is possible within living cells and demonstrated this by using a model of the human telomeric DNA sequence and corresponding RNA template (**Figure 2.11**).[297] It was observed in their study that the CD spectrum of telomeric DNA shows a positive CD signal at around 290 nm and a negative CD signal at around 265 nm (in presence of monovalent Na^+); however, upon addition of the complementary RNA strand to the telomeric DNA, the CD signal at around 290 nm disappeared and a strong positive signal was observed at around 260 nm and a negative signal at around 240 nm that is identical to the transformation of *anti*-parallel DNA to a parallel G-quadruplex. Interestingly the CD spectrum for the hybrid sequence also bears resemblance to the finding of Ren and Chaires. Thus the CD spectrum for the telomeric DNA/RNA hybrid is yet to be established, for which reason, in our study, a control CD spectrum was used that was generated from a specific telomeric DNA/RNA hybrid sequence (*i.e.*, used throughout this project).

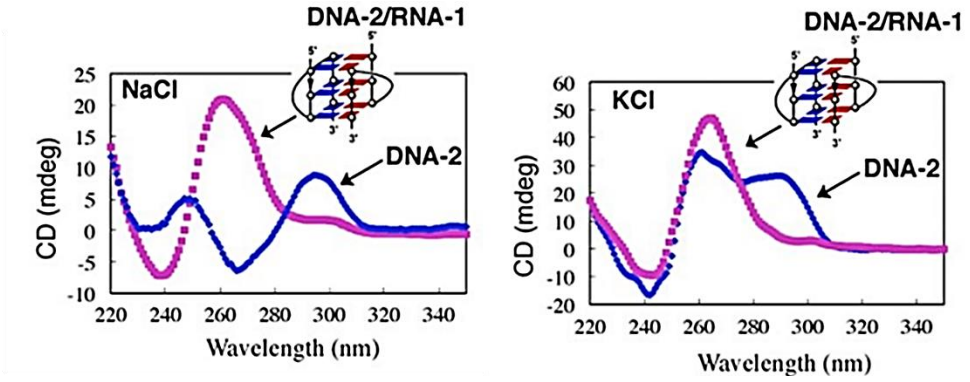
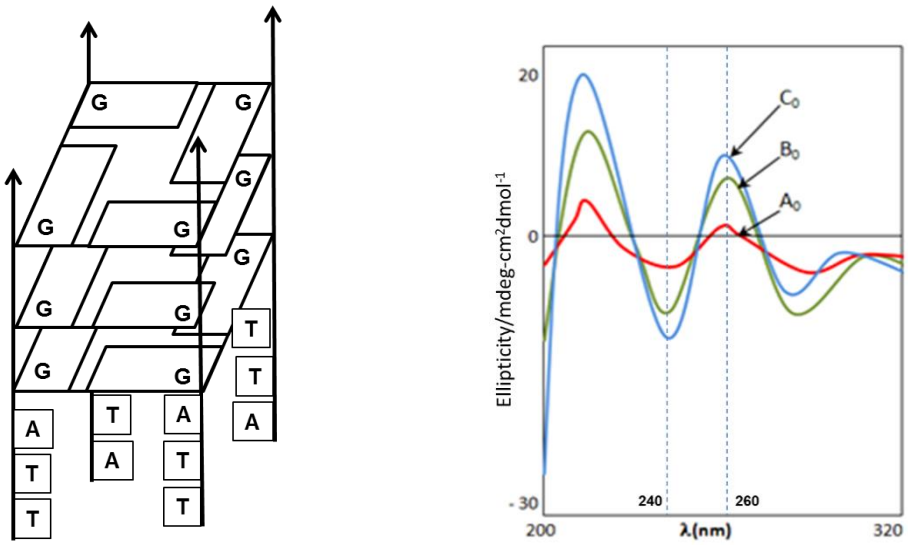


Figure 2.11: Conversion of CD spectra due to formation of hTelo DNA/complementary RNA sequence-assisted hybrid quadruplexes, in presence of both monovalent cation Na^+ and K^+, as reported by Xu and colleagues.[297]

The CD spectroscopy method has become one of the essential techniques for studying G-quadruplex structures, more specifically telomeric G-quadruplexes and ligand interactions. It measures the conformational changes in the DNA structure which are induced by modification of the environment; for example, the nature and/or concentration of counter ions, temperature, pH or the addition of crowding agents as well as chemical modifications.

Human telomeric G-quadruplex DNA (hTelo) can form distinct assemblies due to its special hexamer repeat sequence of TTAGGG. In addition, the conformations formed by hTelo sequences of different lengths are also diverse. Despite having varied spectra due to the length of the repeat sequence, there is some relationship between their CD signals as well. **Figure 2.8** describes the CD spectra of samples A_o, B_o and C_o with mono, di and tetra-repeats of the TTAGGG sequence, respectively; they have a reasonable similarity with a negative CD signal at 240 nm and a positive CD signal at around 260 nm.[298] In this G-quadruplex structure, each strand runs in the same direction, where all guanine residues (from different strands) in each G-quartet plane are in *anti*-conformations.

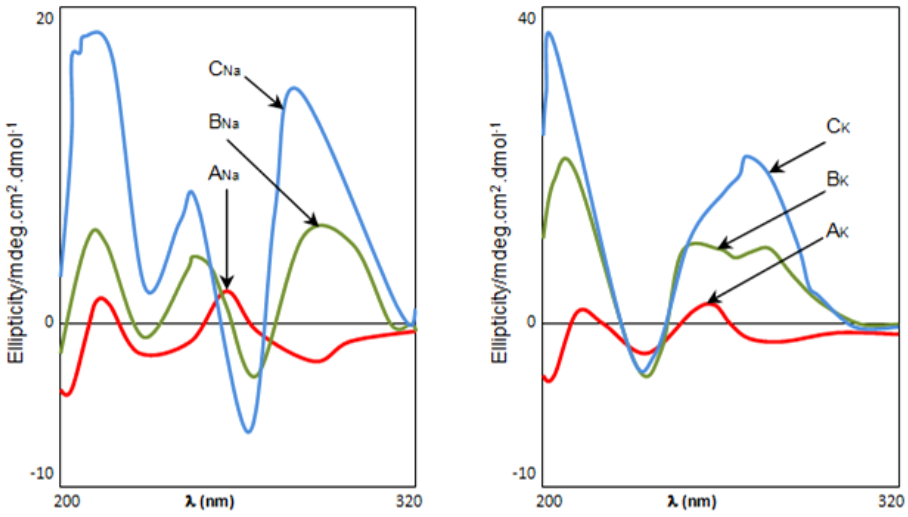


(a) (b)

Figure 2.12: (a) Molecular arrangement and outlined map of parallel quadruplex structure (b) CD spectra reported by Xiaoyan and co-workers for A_o = TTAGGG, B_o = $(TTAGGG)_2$ and C_o = $(TTAGGG)_4$ in Tris-HCl buffer without monovalent metal ions at 20°C, pH 7.5.[299]

CD analysis also showed that there are weak positive signals at around 290 nm in curves of B_o and C_o, where the CD signal for C_o is more intense due to its G-rich sequences that can fold again to form hairpin dimer structures with *anti*-parallel conformation (**Figure 2.8 b**).[300] Contrary to the parallel quadruplex assemblies, the guanine glycosidic twisting angles in the G-quartet can appear in alternative *syn*- and/or *anti*-conformation and thus their CD signal can be characterized by a negative peak at 260 nm and a positive peak at 295 nm.[301]

Additionally, Williamson and co-workers established that addition of monovalent cations (Na^+) changes the G-quadruplex structures remarkably and has different effects on sequences due to varied lengths of DNA repeats.[302] For these reasons, monovalent cations (*e.g.*, K^+ and/or Na^+) are essential for quadruplex formation since they can form a coordination complex to stabilize the G-quartet structure. For example, Xiaoyan and co-workers also reported the effect of Na^+ (80 mrnol/L Na^+ in solution) and K^+ (80 mrnol/L K^+ in solution) on the telomeric G-quadruplex sequence of TTAGGG. The authors observed that due to the variation in telomeric repeat lengths for $(TTAGGG)_n$, the sequence can form an intermolecular quadruplex structure (n = 1), intramolecular hairpin structure and subsequent dimeric *anti*-parallel conformation (n = 2) and monomolecular quadruplex structure (n = 4). However, the aforesaid results have revealed that CD spectra of sequences of different lengths were varied by minor changes. In the presence of 80 mmol/L Na^+ in the buffer solution, A_{Na} (TTAGGG sequence) did not show any obvious change, although, for the CD signals for B_{Na} ($(TTAGGG)_2$ sequence) and C_{Na} ($(TTAGGG)_4$ sequence), at around 240 nm the negative bands disappear, 260 nm positive bands shifted to 250 nm, a new negative signal was detected at around 260 nm and a positive signal was detected at around 290 nm.[299]



(a) (b)

Figure 2.13: CD spectra of TTAGGG (A), $(TTAGGG)_2$ (B) and $(TTAGGG)_4$ (C) in Tris-HCI buffer at 20°C, pH 7.5 in presence of (a) 80 mrnol/L Na^+ in solution, (b) 80 mmol/L K^+ in solution, as reported by Xiaoyan and co-workers.[299]

In 2000, Xiaoyan and co-workers suggested that **i)** the hTelo sequence can form quadruplex structures *in vitro* and in the absence of metal ions can exist mostly in a parallel conformation; however, with the escalation of repetitive frequency of TTAGGG and a subsequent increase in sequence length, hTelo is increasingly likely to form an *anti*-parallel conformation, **ii)** monovalent cations, K^+ and Na^+ can stabilise the quadruplex structures and help in the formation of an *anti*-parallel conformation (note, Na^+ can only stabilise *anti*-parallel conformations whereas K^+ can stabilise both).[299]

2.5.3.1 Procedure involved in CD Analysis

2.5.3.1.1 Oligonucleotides Used in CD Analysis

The 14 selected molecules with good and moderate selectivity towards DNA/RNA hybrid duplexes and hTelo sequences (based on preliminary FRET results and molecular modelling studies) were screened against DRH (DNA/RNA hybrid duplexes), cDRH (control DNA/DNA duplexes), F21T (telomeric quadruplex DNA sequence) (section **2.5.1.1.1**).

2.5.3.1.2 Methods Used for CD Analysis

Simultaneous UV absorption and CD spectra were acquired on either Chirascan or Chirascan Plus spectrometers (Applied Photophysics, Leatherhead, UK). The instruments were flushed with pure nitrogen gas throughout the measurements. Far-UV spectra were recorded from 450 to 200 nm with a 1 nm spectral bandwidth, 0.5 nm stepsize and a 1 s spectrometer time-per-point. A rectangular 0.5 mm path length was employed. Unless otherwise stated, all spectra were measured at 25 °C (temperature controller made by Quantum NorthWest, Model-TC125). During data processing, a spectrum of the DNA media or solution was buffer-subtracted and Savitsky-Golay smoothing with a convolution width of 5 points applied. CD spectra were normalized for concentration and path length and expressed in terms of molar ellipticity per residue. Secondary structure analyses were performed using CDPro.

Preparation of CD-Buffer: Two types of buffer were prepared with Tris (Tris(hydroxymethyl)aminomethane), hydrochloric acid (HCl) and sodium chloride. For DRH and cDRH sequences, 50 mM Tris was prepared, adjusted to pH 7.4 with hydrochloric acid (1 N) and labelled as 50 mM 'Tris-HCl'. For the F21T sequence, 200 mL of buffer solution was prepared with 50 mM Tris and 100 mM sodium chloride in distilled water, adjusted to pH 7.4 with hydrochloric acid (1 N) and labelled as 'Tris-NaCl-HCl'. The buffer was stored in the freezer at -20 °C.

Preparation of DNA: The oligonucleotide sequences were diluted with sterile DEPC water (DNA Grade, Fisher Scientific) to obtain 100 µM solutions. Five µM solutions of DNA were prepared using CD-buffer (Tris-HCl for DRH and cDRH, Tris-KCl-HCl for F21T). The DNA solutions were stored in the freezer at -20 °C.

Preparation of Ligand Solutions: For all 14 ligands, 5 mM stock solutions were made with dimethyl sulfoxide (≥ 99.9 %, A.C.S. spectrophotometric grade, Sigma-Aldrich). From these stock solutions, 5 µM, 10 µM, 15 µM, 20 µM and 25 µM working solutions were prepared during experimentation with DNA solutions.

Preparation of Sample for Binding Titration Curve: For the CD-based DNA binding titration assay, 1,000 µL of respective buffer was first scanned and subsequently 1 µL of ligand was added to the same cuvette (filled with buffer) to scan one more time. Following the first step,

the cuvette was cleaned properly and dried with a N_2 flow. Afterwards, 1,000 µL of DNA solution was scanned to measure the appropriate wavelength for each type of DNA sequence. After determining the proper wavelength, 1 µL of ligand solution was added to the same cuvette to make a ligand concentration of 5 µM, which was then scanned again. Subsequently, an additional 1 µL of ligand solution was added into the cuvette to make the 10 to 25 µM ligand concentrations for each scanning. CD measurements were taken at 25 °C at a wavelength of 200-450 nm.

2.5.3.2 Limitations of CD spectroscopic analysis

Measurement of circular dichroism is difficult in the presence of typical aqueous buffer systems that can show differential absorption of circularly polarized light. That is why phosphate, sulphate, carbonate, and acetate buffers are generally used, if necessary, as extremely dilute solutions (*e.g.*, 10-50 mM). In the case of far-UV CD analysis, a TRIS buffer system is completely avoided. In some cases, fluoride needs to be substituted for chloride ion as fluoride absorbs less in the far-UV. To minimize solvent absorption, shorter path length cells (*e.g.*, 0.1 mm) are used for the far-UV CD analysis. In addition, for π to π^* orbital absorption bands of the amide bonds lie partly in the vacuum ultraviolet region (*i.e.*, wavelengths less than about 200 nm), which is essentially inaccessible in air because of the strong absorption of light by oxygen at these wavelengths. To eliminate this problem, a continuous nitrogen flow is supplied to the instrument to confirm an oxygen-free instrument.

2.5.4 Cytotoxicity Test with MTT Assay

Cytotoxicity is the feature of any molecule or drug showing toxicity and/or growth-inhibitory effect towards a targeted cell line. In research, cytotoxicity studies are broadly carried out to screen molecules for their cytotoxicity. It is a balancing tool to ensure a novel molecule is toxic to the desired cell line (for example, breast cancer cell line, MDA-MB-231) and safe for normal healthy cell lines, so as to avoid an adverse toxic effect. Researchers usually look for cytotoxic compounds to develop therapeutic agents aiming for the quickly-dividing melanoma cells or screen ‘hit molecules’ from preliminary high-throughput drug screenings for undesirable cytotoxic effects.

Cytotoxicity can be evaluated by using the MTT dye (3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2*H*-tetrazolium bromide), for which the assay is popularly known as the MTT assay. A colorimetric reaction is used in this assay to measure the reducing potential of the cell. In this assay, usually viable cells reduce the MTT reagent (**2.3**) to a coloured formazan product (**2.4**), which is visible under a photometer (usually between 500 and 600 nm). There are several tetrazolium dyes, like XTT, MTS and WSTs, which are also used in conjunction with the intermediate electron acceptor 1-methoxy phenazine methosulphate (PMS) for the cytotoxicity assay. The sulforhodamine B (SRB) assay is another cytotoxicity assay that is widely used by different research groups.

The MTT assay relies on cell metabolic activity, where the MTT is reduced as a result of NADH formation by the metabolically active cells and thus NADP(H)-dependent cellular oxidoreductase enzymes reflect the number of viable cells present under defined conditions (**Figure 2.14**). These enzymes are capable of reducing the tetrazolium dye MTT (**2.3**) to the insoluble, purple-coloured formazan (**2.4**).[303, 304] Usually, the MTT assay is carried out in dark conditions due to its light sensitiveness. As the NAD(H) is mostly formed within the mitochondria of the cell, 1-methoxy-phenazine methosulphate (**2.5**) is needed as the electron-carrier to transfer the reduction equivalents from the mitochondial matrix into the medium. Due to its mechanism of action, the MTT assay can only detect the viable cells. Tetrazolium dye assays are used to evaluate the cytotoxicity (loss of living cells) or the cytostatic efficacy (modification from proliferation to latency of active cells) of prospective drug-like molecules and toxic materials.

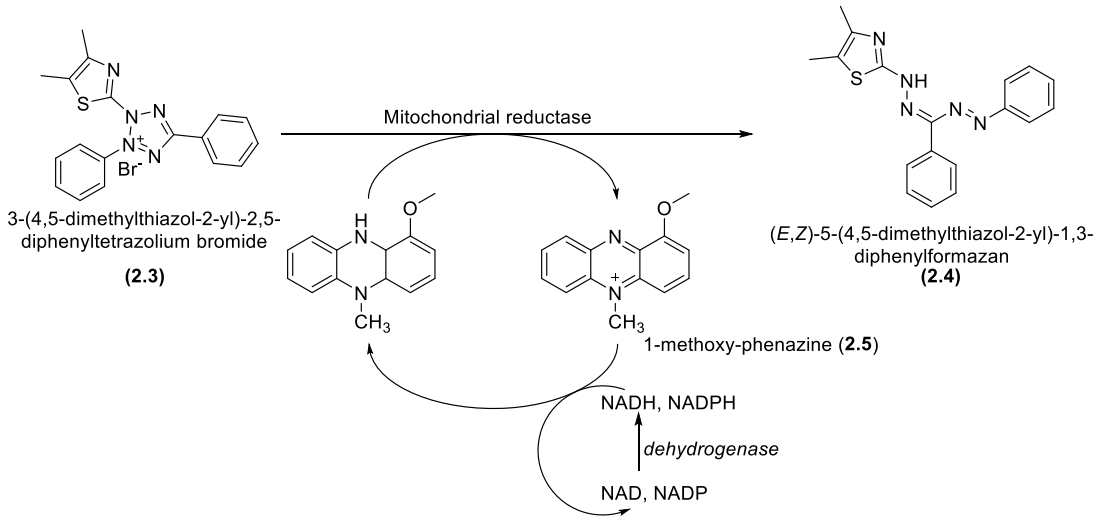


Figure 2.14: Schematic diagram illustrating the conversion of MTT to formazan (purple) in the presence of an electron acceptor, 1-methoxy-phenazine.

2.5.4.1 Procedure involved in MTT Assay

2.5.4.1.1 Cell lines used for MTT Assay

The HeLa (Cervical cancer), MDA-MB-231 (Triple negative breast cancer) and NCI H1975 (Non-small cell lung cancer) cell lines used in this assay were obtained from the American Type Culture Collection (Manassas, VA).

2.5.4.1.2 Method and Materials for MTT Assay

The MTT assay is comprised of a number of crucial steps, from cell splitting to 96-well plate reading in a photometer. The steps below were followed to process the MTT assay for each type of cell line:

Cell Culture: HeLa and MDA-MB231 cell lines were cultivated in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% foetal bovine serum (FBS), 1% penicillin and streptomycin and 1% MEMNEAA (200x) non-essential amino acids. NCI H1975 cell lines were grown in RPMI-1640 supplemented with 10% foetal bovine serum (FBS) and 1% penicillin and streptomycin. These cells were maintained in an incubator at 37 °C with 5% CO_2.

Cell Passaging: Cell passaging or splitting is a procedure that allows cells to be kept alive and growing under culture conditions for prolonged periods. Cells should be passed when they are 90-100% confluent. In this process, a small number of confluent cells are transferred into

a new vessel since a high density of rapidly dividing cells is related to cells going into senescence. Suspension cultures can be simply split with a small amount of culture comprising a few cells diluted into a larger volume of fresh media. However, in the case of adherent cultures, cells first need to be detached with trypsin, and then a small number of detached cells can be used to seed a new culture.

Ethanol (70%) was used to clean the fume hood, equipment, flask and pipette before starting any process. By that time, DPBS (Dulbecco's Phosphate Buffered Saline) and DMEM (Dulbecco's Modified Eagle Medium) were taken out of the fridge to become warm and ready to use. Cells contained in a T-flask were taken from the incubator and checked under a microscope (dead cells floated and appeared round-shaped, whereas growing cells were attached to the bottom surface in a rod-shaped and clustered fashion). Previous media was aspirated by using a vacuum pump (a sharp pointed pipette, 2 mL, was used for this purpose). Five mL of DPBS was added to wash previous media fully and clean the T-flask, which was aspirated again to remove the floating dead cells as well. One mL of trypsin-EDTA was added to detach the growing cells and the flask was then incubated at 37 °C in a 5% CO_2 incubator for 2 minutes. Ten mL of DMEM was added to inactivate the trypsin; afterwards the total suspension was transferred into a falcon tube for centrifugation (2.5 minutes, 21 °C, 1.5 rpm). All the supernatant medium was aspirated out, leaving the cells at the bottom of the tube. Finally, 5 mL of DMEM was added and mixed properly and an appropriate amount of cell suspension transferred to a newly-labelled T-flask according to the next passage time. Furthermore, an appropriate amount of DMEM was added to the same T-flask to make a total volume of 15 mL (in general, 15 mL is the optimum volume for the survival of cells in a T-flask). Lastly, the new T-flask was returned to the incubator (37 °C, 5% CO_2) for the next splitting schedule. Cells were passaged every 2 to 3 days.

Cell Count and Plate Preparation (Seeding of Cells): Seeding is the process of spreading a defined amount (in volume or cell number) of a cell suspension onto a plate. For our experiments, 96-well polypropylene plates were used which are free of binding affinity for proteins or DNA, allowing complete sample recovery. These plates can withstand temperatures of -80 to +121°C. For the seeding process, initially general cell culture procedure

and precautions were taken (*e.g.*, 70% ethanol was used to clean the fume hood, equipment, flask and pipette, DPBS and DMEM were taken from the fridge to become warm and ready to use, cells contained in the T-flask were taken out from the incubator and checked under a microscope). Afterwards previous media was aspirated by using a vacuum pump, and 5 mL DPBS was added to wash previous media fully and clean the T-flask, which was aspirated again to remove the floating dead cells as well. One mL trypsin-EDTA was added to detach the growing cells and the flask was incubated at 37 °C in 5% CO_2 incubator for 2 minutes. Ten mL DMEM was added to inactivate trypsin; afterwards the total suspension was transferred into a falcon tube for centrifugation (2.5 minute, 21 °C, 1.5 rpm) and all the supernatant medium was aspirated out leaving the cells at the bottom of the tube. Then, 1 mL of DMEM was added to suspend the cells. A haemocytometer was used to count cells, where a mixture of cells was made with typan blue stain (10 µL cells/90 µL stain) to visualize the cells under an electron microscope. An appropriate amount of cell suspension was taken to make a dilution of cells with DMEM to confirm about 10^6 cells per well. Afterwards, a multichannel pipette was used to seed cells onto the plate, which was then returned for incubation at 37 °C in a 5% CO_2 incubator (incubation time was varied from 24-72 hours).

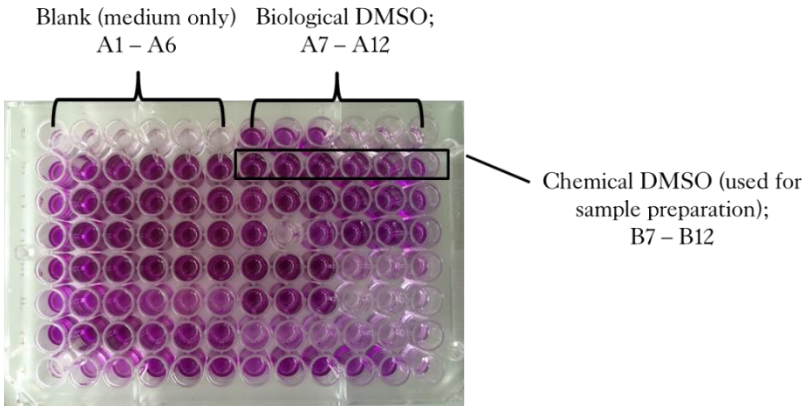


Figure 2.15: MTT assay plate layout in general.

Addition of Tested Sample: In this stage, the previous medium was aspirated out from every well by using sharp pipette tips and replaced with an equal amount (100 µL) of fresh medium and ligand of appropriate concentration, made from a 5 mM stock solution by using DMSO, which were then added onto the cell seeded plate. Lastly, the plate was returned for further incubation at 37 °C in a 5% CO_2 incubator (incubation time was varied from 24-72 hours).

Plate Reading and Spectrometric Analysis: The final stage was reading the data from the processed plate. Firstly, old medium was aspirated off from each well after the predetermined incubation period and subsequently each well was washed with 100 µL of medium (high glucose content but without phenol red). Then all medium was aspirated off from each well and 100 µL of previously made MTT/medium phenol red-free solution was added into each well. A further 4 hours of incubation was undertaken at 37 °C in a 5% CO_2 incubator and afterwards the medium was aspirated off from each well and 100 µL of DMSO added to dissolve the crystals formed. The plate was again incubated for 5 minutes at 37 °C in a 5% CO_2 incubator and subsequently placed in a shaker for 5 minutes (500 rpm) to remove all air bubbles.

Absorbance was then taken by an Infinite 200Pro plate reader at a 570 nm wavelength and the data were processed with the help of Tecan i-control application software.

Of note, MTT solution was prepared in a 1:10 dilution with phenol red-free medium; *e.g.*, 1 mL of MTT was added into 9 mL of phenol red-free medium. The final solution was filtered through a 0.2 µM filter and kept in the dark at 4 °C. Such solutions can be stored for up to a month.

2.5.4.2 Limitations of MTT assay

MTT assay has many limitations such as lack of sensitivity, chemical interference, toxicity, and batch-to-batch variation of reagents. The absorbance method of detection used by the MTT assay is generally less sensitive than fluorescent and luminescent methods for spotting viable cell number. The detection sensitivity differs widely among cell types and depends on the metabolic activity of the tested cell type. Usually tetrazolium reduction assays can detect 200-1000 cells per well under ideal settings which can be improved by optimizing the concentration of MTT and the incubation time; although the incubation time is restricted by the toxicity of the MTT reagent. Assay conditions affect the chemical or enzymatic reduction of MTT. A variety of chemical compounds (*e.g.*, reducing agents like ascorbic acid, glutathione, coenzyme A, polyphenolic compounds) are known to interfere with the MTT assay. Moreover, chemicals that uncouple electron transport from oxidative phosphorylation of ATP can interfere with the MTT assay. Continuing exposure of MTT reagent to light and high pH of the culture medium can also result in production of formazan and higher background absorbance. Appropriate

controls are necessary to avoid such chemical interferences. Riss and co-workers also reported that MTT reagent exhibits cytotoxic effects to eukaryotic cells.

Although the MTT assay results associate with the number of viable cells in standard culture conditions, the rate of tetrazolium reduction may imitate the general metabolic activity or the rate of glycolytic NADH production. The rate of MTT reduction can change with pH, glucose content of medium and the physiological state of the cells. Mitochondria are not the only source of MTT reduction; more recent studies revealed that NADH is responsible for most MTT reduction which is associated with the cytoplasm, membranes in the endosome/lysosome compartment and the plasma membrane, along with the mitochondria. It is now taken into consideration that MTT reduction at the plasma membrane is responsible for the observation of formazan crystals taking place outside of cells.[305, 306]

2.5.5 Preliminary Biological Evaluation

2.5.5.1 Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) Assay

Most of the methods used for studying DNA/ligand interactions are based on the polymerase chain reaction, or PCR, and one of the most commonly used methods in PCR is RT-PCR. This is a useful tool to study the interference of drug molecules with polymerase activity to obtain evidence about ligand-mediated stabilization of a targeted DNA sequence (for example, a G-quadruplex DNA sequence).[307]

RT-PCR is a qualitative measure of gene expression that works through the formation of complementary DNA (cDNA) transcripts from RNA using reverse transcriptase enzyme. Then, the resulting cDNA is amplified with the help of the polymerase chain reaction.[308] The PCR procedure involves denaturation, annealing and extension reactions of the oligonucleotide to achieve the amplified PCR product.[309] As RT-PCR is a qualitative measure for gene expression, it is a useful tool for detecting the mRNA expression.[310]

In order to determine the gene expression for a tumour or melanoma cell growth, the RNA is extracted from the targeted cancer cells and subsequent synthesis of cDNA is done with the help of a reverse transcriptase enzyme. The desired gene and internal standards are then intensified using the PCR method in a separate reaction tube. Afterwards, the gel

electrophoresis, autoradiography and detection processes are carried out sequentially in later steps. Lastly, the relative measurement of gene expression is made according to the up-regulation or down-regulation of expressed genes and by comparing the size of the bands (spots) of sample and control in gel electrophoresis.[240]

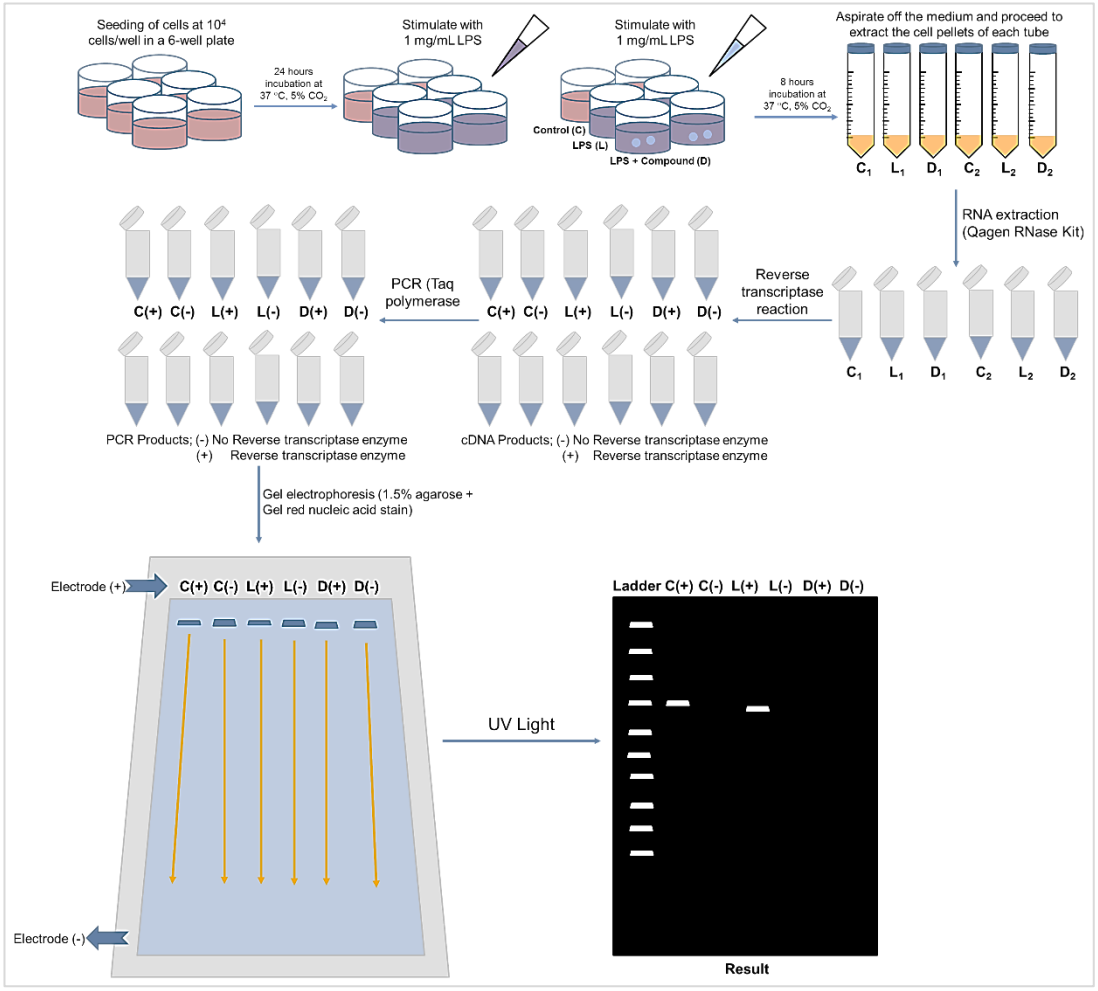


Figure 2.16: Schematic representation of the steps involved in the RT-PCR process.

To quantify the gene expression, real-time RT-PCR is also extensively used due to its quick screening capabilities for large numbers of samples and can be implemented with the Light Cycler hTERT.[309, 311] Emrich and co-workers describe RT-PCR as a rapid method that combines amplification and detection in a single step and as a sensitive technique which is advantageous for a limited amount of sample.[309]

Telomerase activity has a significant relationship with the expression of human telomerase reverse transcriptase (hTERT).[308, 312] Thus this is a useful tool to evaluate any novel drug/small molecule for its telomerase inhibitory efficacy. It is to be noted that hTERT mRNA is overexpressed or upregulated in some tumours, for example ovarian malignant tumours[312] and hepatocellular carcinoma[313]. As the activity of the hTERT promoter is suppressed in most human somatic cells, hTERT is an appropriate indicator to use to characterise the expression of cytotoxic genes in malignant cells.[311] In 2007, Sun and co-workers reported the application of RT-PCR to investigate telomerase activity. The authors also suggested that hTERT expression is the rate-limiting factor of telomerase activity and that it could be a possible marker for ovarian cancer investigation.[312] In 2011, Shay and co-workers established that treatment with cisplatin could downregulate telomerase activity and hTERT mRNA expression in a time- and dose-dependent manner.[314] Thus both studies involving RT-PCR showed a significant connection between hTERT mRNA expression and telomerase activity in telomerase-positive cells.

2.5.5.1.1 Procedure involved in RT-PCR Assay

The MDA-MB-231 cell line was used in the RT-PCR assays. The cells were seeded at a cell density of 1 x 10^6 cells per well in a 6-well plate. The plate was incubated for 24 hours for cell attachment. After that, cells were stimulated with 1 mg/mL *Escherichia coli* 026:B6-LPS (Sigma) in LPS and compound wells. The compounds at concentrations of either 1 µM or 10 µM were used in the cell stimulation and were incubated for 8 hours. Cell pellets of MDA-MB-231 cells from each well (control, LPS, and LPS plus compound wells) were extracted. After cultivation, an RNeasy kit (Qiagen) was used to extract RNA from the cell pellets. RNA was reversed transcribed to a cDNA template using a High Capacity RNA-to-cDNA Kit (Applied Biosystems). The amplification of hTERT and cyclin B genes was performed with Platinum Taq DNA polymerase (Invitrogen). The primer sequence used in this study for hTERT was 5′-GGAGCAAGTTGCAAAGCATTG-3′ (F) and 5′-TCCCACGACGTAGTCCATGTT-3′ (R).[315] The primer sequence of cyclin B was 5′-AAGAGCTTTAACTTTGGTCTGGG-3′ (F) and 5′-CTTTGTAAGTCCTTGATTTACCATG-3′.[316] PCR amplification was performed as follows; 30 cycles of denaturation at 95 °C for 30 sec, annealing at 55 °C for 30 sec and extension stage

at 72 °C for 1 minute 30 sec. The amplified products were analysed by gel electrophoresis using 1.5% agarose gel (Sigma) and stained with GelredTM nucleic acid gel stain 10,000 X (Biotium). The product band was captured under ultraviolet light. The band intensity of sample was compared with a control.

2.5.5.2 Electrophoretic Mobility Shift Assay (EMSA)

In general, the electrophoretic mobility shift assay (EMSA) is an affinity electrophoresis procedure by which protein/DNA or protein/RNA interactions can be studied. The assay is usually carried out on either a polyacrylamide (PAM) gel or agarose gel for a short period of 1.5-2 h for a 15- to 20-cm gel.[317] Upon application of a low electric field, the macromolecules (for example, DNA, RNA, protein) themselves and/or the drug-macromolecule conjugates move through the gel at a variety of speeds due to their different sizes and charges. In practice, the control lane (the macromolecule without any ligand or drug) contains a single band corresponding to the unbound oligonucleotide fragment. However, if the ligand or drug is capable of binding to the oligonucleotide or other macromolecules, the lane with drug (or compound) will show a larger band shifted up in the second ladder, ultimately indicating a positive sign of drug-DNA binding.

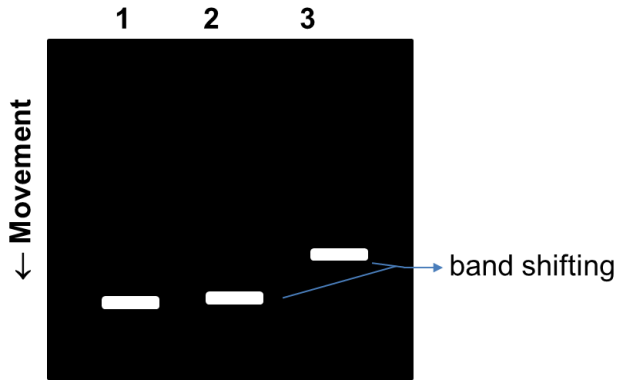


Figure 2.17: The general principle of the electrophoretic mobility shift assay.

Though the electrophoretic mobility shift assay was developed to study the DNA-binding properties of proteins, in this project the assay was primarily used to investigate the shifting of ligand-bound DNA compared with DNA itself.

2.5.5.2.1 Procedure involved in EMSA

Electrophoretic mobility shift assays were performed as binding analyses. Band shifts were detected under ultraviolet light. One hundred μM of the oligonucleotide DNA/RNA hybrid duplex (DRH, TTAGGGTTAGGGTTTTTTGGG) was used as a control. The experiment was performed in 4% agarose gel. The compounds were tested at a 2.5 μM concentration diluted with 10-times Tris Acetate EDTA (TAE) buffer. Loading conditions were as follows: the control well was a mixture of 10 μL of Oligo-DRH and 10 μL of TAE buffer solutions, and test wells were mixtures of 10 μL of Oligo-DRH and 10 μL of compound (2.5 μM) solutions. Bands were photographed under ultraviolet light.

Chapter 3A: Chemistry: Results and Discussion

3A.1 Introduction

The DNA/RNA hybrid duplex is a unique structure within the human body which typically forms during the catalytic activity of the upregulated telomerase of malignant cells. In order for telomerase to function correctly, this heteroduplex assembly needs to go through subsequent association and dissociation of the RNA template from the telomeric overhang. Thus it is hypothesised that stabilization of the heteroduplex after association steps can stop the dissociation of the RNA template and successively allow the DNA to signal entry into senescence. Though this DNA/RNA hybrid duplex may be a unique and specific target, there were no previously reported works found concerning it and relating to the field of drug discovery. In 2010, Wheelhouse and co-workers reported 4,6-*bis*-[4'-[[2''-(dimethylamino)ethyl]-mercapto]phenyl]pyrimidine dihydrobromide as a preferential ligand for telomeric quadruplexes, triplex DNA structures and hybrid poly(dA)·poly(rU) heteroduplexes (**1.23**, **Figure 1.37**).[192] Though the heteroduplex structure used in their study was not a telomeric DNA/RNA hybrid duplex, the authors demonstrated a *bis*-molecule accompanied by a tertiary amine tail that can interact and/or wrap around the minor groove of the DNA/RNA hybrid duplex. It has been confirmed that the molecule has an approximately 20-fold selection preference for both backbone and base orientation in the poly(dA)·poly(rU) heteroduplex over the comparable RNA, 7-fold over the alternative poly(rA) and 3-fold over duplex DNA.

However, the study of Wheelhouse and co-workers in 2010 concluded that a G-quadruplex binder can also stabilize the DNA/RNA hybrid structure as well. Assuming this assertion to be true, a number of molecules were reviewed in this project to identify a suitable lead for further developments of new analogues which could be drug candidates.

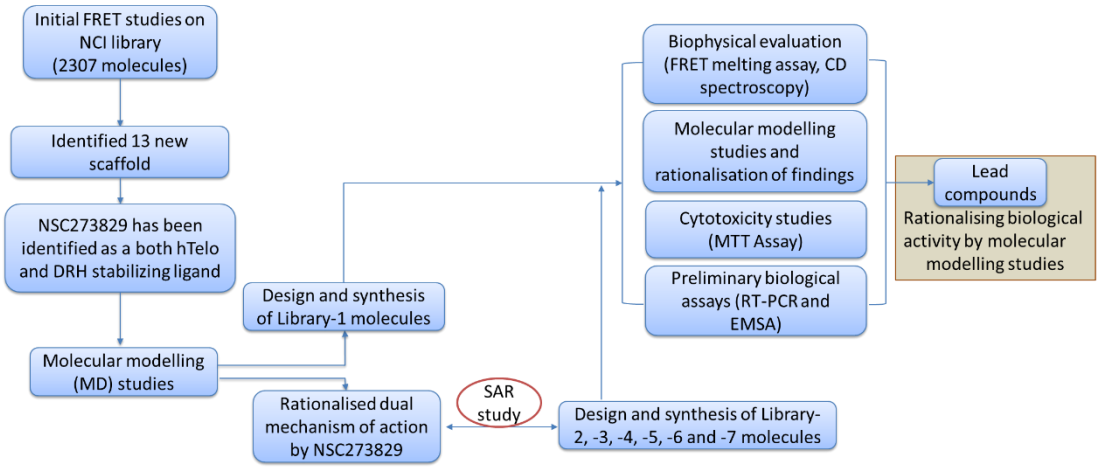


Figure 3A.1: Experimental approach to develop small molecule targeting telomeric DNA/RNA hybrid duplex.

In 2012, an article based on screening 2307 NCI library molecules by Rahman and co-workers provided new chemical scaffolds with which to develop G-quadruplex targeting agents.[261] Based on preliminary FRET melting results and previously reported key biological activity, they identified 13 different scaffolds (for example, *bis*-benzothiazole, tetrahydropyrimidine, hydroxystilbamidine, *bis*-dihydro-1H-imidazolphenylisophthalamide, *bis*-quinoline-amine, hydroxyquinoline-quinolinium, quinoline-amine-acrylamide and fused imidazo-quinoline scaffolds) from the NCI Diversity Set II, the NCI Natural Products Set II and the NCI Mechanistic Diversity Set libraries. Among these 13 scaffolds, compound NSC 273826 showed a 20-fold binding preference towards telomeric G-quadruplex over a duplex DNA sequence. Additionally, the molecule was reported as active against L1210 leukaemia xenografts (in mice).[262] This *bis*-quinoline-amine ligand with a heptamethylene spacer (NSC 273829) also showed about 3.2 °C stabilization towards DNA/RNA hybrid duplex structures (in-house unpublished data). As the compound falls into the Lipinski Rule of Five for drug-likeness[318], this project was developed to modify the compound and synthesize and/or rationalize the importance of the spacer length in stabilising the DNA/RNA hybrid duplex.

To optimise the identified hit compound NSC 273829, a series of analogues were synthesized with the help of a molecular modelling study and solution phase chemistry. Initial molecular modelling suggested incorporation of a methylene spacer between the *bis*-intercalating group

and the terminal phenyl group that were connected by methylene spacers containing 4-10 carbons. The following general structure was identified as the early synthetic candidate for this project.

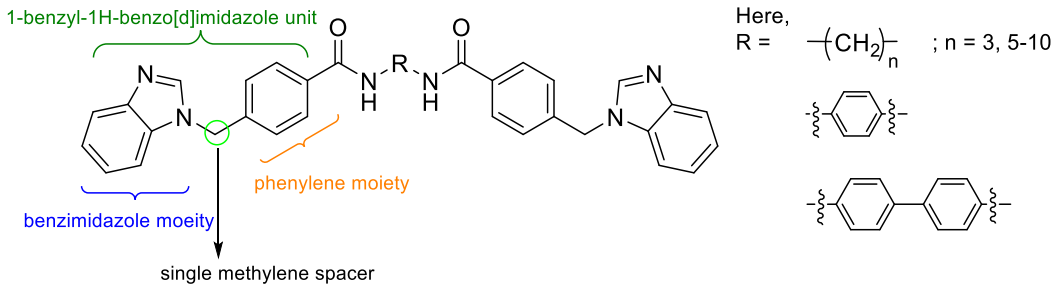


Figure 3A.2: General structure of the library-1 molecules.

To explore further the role of methylene linkers and the terminal *bis*-intercalating group, a number of focused libraries were synthesized. The general structures of these libraries are summarised below.

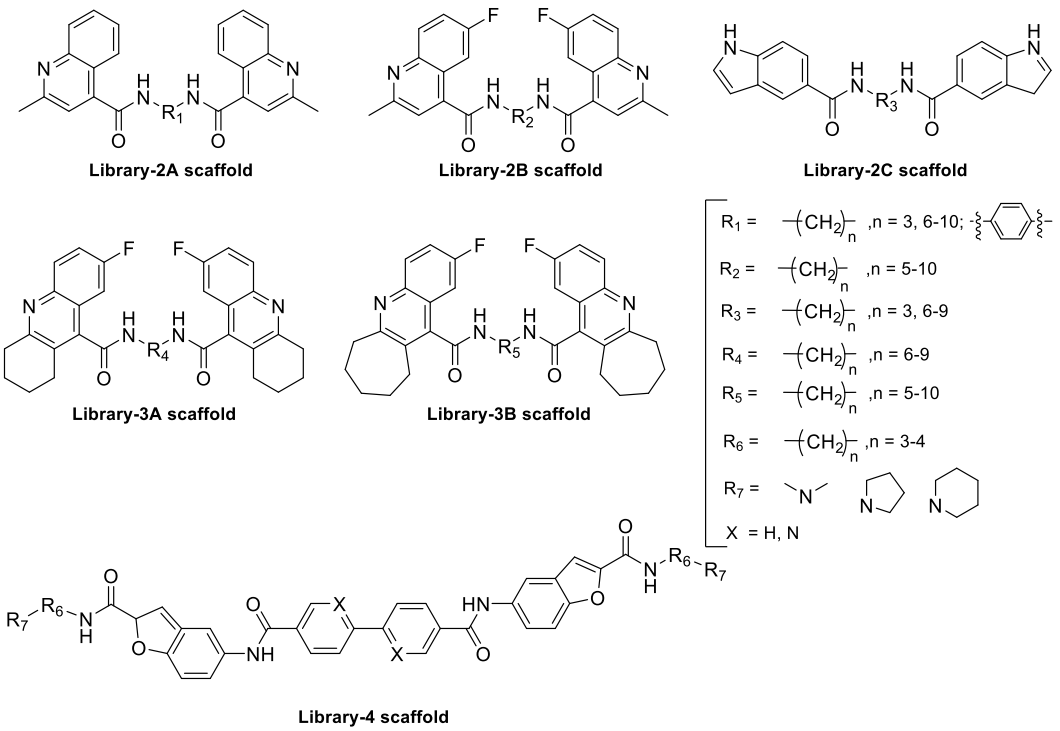


Figure 3A.3: General structures of other library molecules synthesized in this project.

3A.2 Molecules in Library-1

3A.2.1 Design Rationale of Library-1 Molecules

In the NCI library molecule NSC 273829, two 7-amino-2-methyl quinoline side arms were connected by a 7-methylene spacer in secondary amine form (**Figure 3A.4**). This project was designed to synthesize more analogues of NSC 273829 with increased binding affinity and selectivity towards DNA/RNA hybrid duplexes. In an initial attempt, molecular modelling was used where it was identified that the 1-benzyl-1*H*-benzo[*d*]imidazole moiety can fit into the telomeric DNA/RNA hybrid duplex. It was also confirmed by the molecular modelling studies that the single methylene spacer between benzimidazole and phenyl moieties is responsible for a kink the structure and the pointing of the nucleobases into the DRH duplex structure (**Figure 3A.4**).

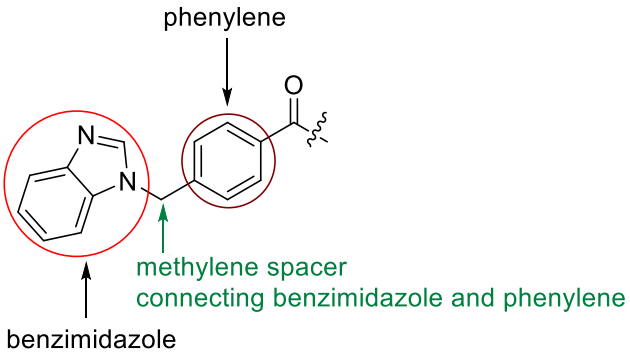


Figure 3A.4: The benzimidazole building block used in the synthesis of library-1 molecules.

In addition, benzimidazole was used as side bulky arms with the aim of being the intercalating part of the molecule. It is already known that heterocyclic compounds are very commonly utilized in the development of new drugs with new modes of action. Among heterocyclic scaffolds, benzimidazole products have a very significant role due to their wide spectra of biological functionalities. A benzimidazole core and its analogue compounds were described as being physiologically and pharmacologically potent and it was observed that an application of this group of pharmacophores can be used for treating some diseases like epilepsy, diabetes and infertility. Additionally, benzimidazole core-containing molecules have antibacterial, antifungal, antitubercular[319], antimalarial[320], anti-inflammatory, analgesic,

antiamoebic[321], antihypertensive[322], antiallergic[323], antikinase and anti-HIV-1[324], antiulcerative, antioxidant and anticancer[233, 325], antiproliferative [232] and antitumor[326-328] activities. With the help of molecular modelling, benzimidazole connected with a phenylene moiety by a single methylene spacer was chosen for synthesis of the first set of molecules in library-1.

To investigate the importance of the length of the methylene spacer between the two terminal building blocks, varied numbers of methylene linkers were utilized while keeping them relevant to the lead compound, NSC 273829. Moreover, phenylene and biphenylene spacers were also used in synthesizing library-1 molecules to evaluate the impact of rigidity on compounds targeting DRH sequences.

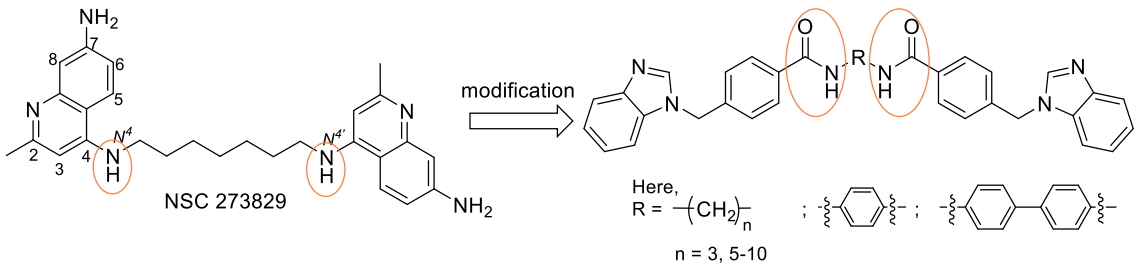


Figure 3A.5: Comparison between library-1 molecules and NSC 273829.

3A.2.2 Synthesis of *bis*-benzimidazole molecules in library-1

The general HOBt/DIC-mediated amide coupling method, as discussed in section 2.4, was employed to prepare the library-1 molecules, which are shown in **Table 3A.1**. All the compounds were synthesized at room temperature using dimethylformamide as a solvent (general reaction scheme is shown in **Figure 3A.6**) and the compounds were labelled as *bis*-benzimidazole molecules. Most of the reactions were straightforward and were finished after being left to react overnight.

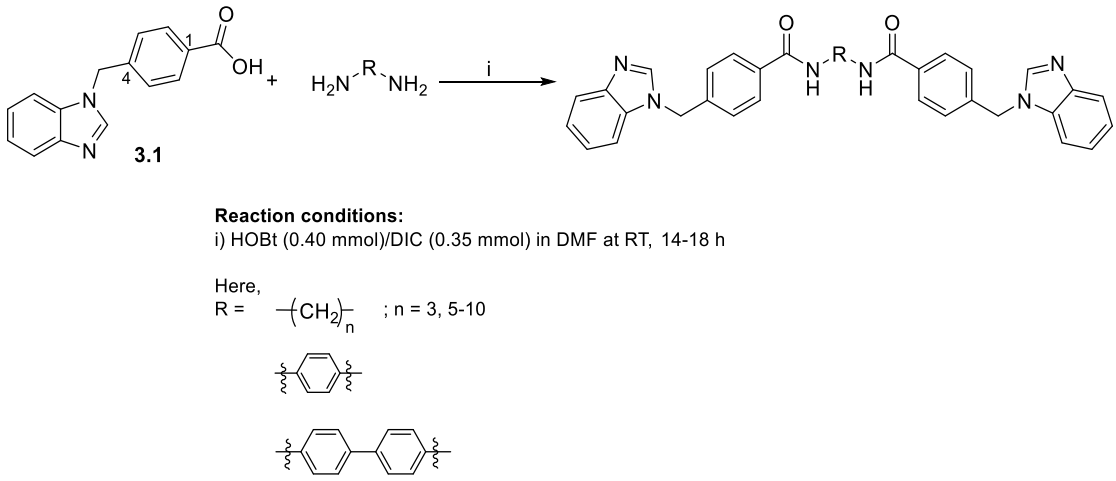
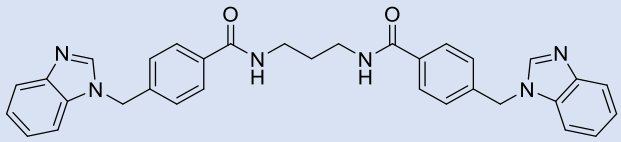
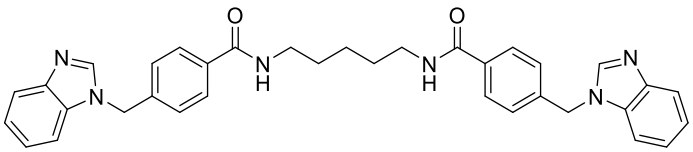


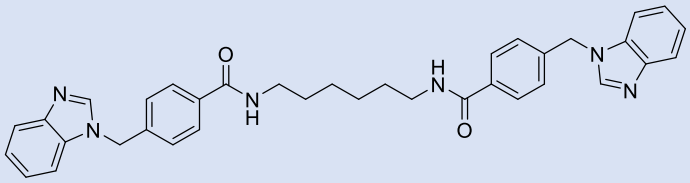
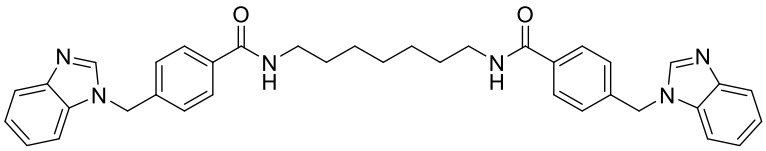
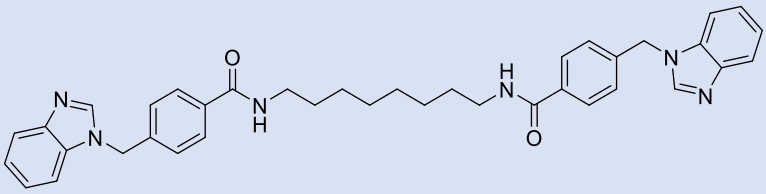
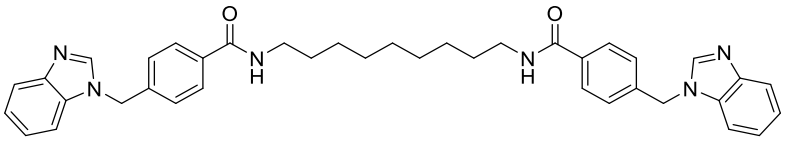
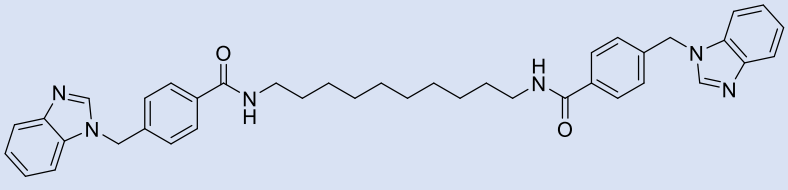
Figure 3A.6: General synthetic scheme for library-1 molecules.

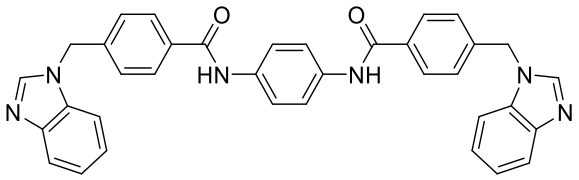
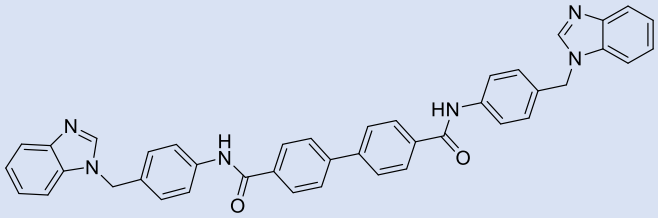
In general, to synthesize the library-1 molecules a solution of 4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzoic acid (0.20 mmol) in dimethylformamide (7 mL) was treated with 1-hydroxybenzotriazole (HOBt) (0.40 mmol) and *N,N'*-diisopropylcarbodiimide (DIC) (0.35 mmol) at room temperature. After an initial activation step (formation of activated acid ester, usually 20-30 minutes), the respective diamine (0.24 mmol) was added to the reaction mixture and allowed to stir overnight. To achieve compound **3.5**, after initial observation (approximately 9 h) *via* LC-MS, an additional 0.10 mmol of acid was added to consume excess amine. Upon confirmation of bi-adduct formation *via* LC-MS after another 7 h, the reaction was quenched with water and extracted with ethyl acetate (discussed in experimental section, Appendix 1). Compounds **3.6** and **3.7** were also achieved *via* a similar process; however, compound **3.8** could not be synthesised in this way. During quenching of the reaction with water, the reaction mixture formed a cloudy solution. While ethyl acetate was being used to extract the organic phase, a thick white precipitate was formed. This was collected by filtration and freeze dried to achieve the compound. Compounds **3.2** and **3.9** were also achieved by following a similar protocol to that used for compound **3.8**. In the case of the synthesis and purification of compound **3.4**, a similar protocol was used; however, neither the desired product's NMR spectrum nor its mass-to-charge ratio (m/z) were observed in the ethyl acetate fractions. Hence it was decided to extract the compound by using methanol instead of ethyl acetate and finally the product was collected in the methanol extract but in lower amount than expected (yield 52%). As methanol is miscible with water (used for quenching the reaction), it was assumed that a portion of the compound was left in the water phase. Compound **3.3** could not

be purified straight away using the extraction method. Due to the presence of impurities in the combined ethyl acetate fractions, flash column chromatography was employed to achieve the pure compound **3.3**. However, due to a second round of purification, a good amount of compound **3.3** was either lost in the mobile phase or stacked in the stationary phase and ultimately gave a yield of 23%. Synthesis of compound **3.10** was difficult, and in the initial ~24 h of reaction LC-MS indicated that there was no formation of product. Hence an additional 0.20 mmol of acid was added to the reaction mixture and left for another ~24 h to form the complete product. Almost two days were needed to consume all the reactant diamine and form the desired product, **3.10**. Upon the addition of water (to quench the reaction), a white thick precipitate was formed (similar to compound **3.8**). Ethyl acetate was added, but the whole reaction mixture became thicker with the addition. Afterwards, n-hexane was used in the reaction mixture, which allowed the total mixture to be cooled in an ice bath. Finally, the product precipitated from the reaction mixture, and was removed by filtration freeze dried to achieve the pure compound.

Table 3A.1: *Bis*-benzimidazole molecules in library-1.

Compound	Structure (Chemical name, formula, MW, log*P*)	Yield (%)
3.2	 *N*,*N*'-(propane-1,3-diyl)bis(4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzamide) Chemical Formula: $C_{33}H_{30}N_6O_2$ Molecular Weight: 542.64 Log P: 3.7	29
3.3	 *N*,*N*'-(pentane-1,5-diyl)bis(4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzamide) Chemical Formula: $C_{35}H_{34}N_6O_2$ Molecular Weight: 570.70 Log P: 4.58	23

3.4	 *N*,*N*'-(hexane-1,6-diyl)bis(4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzamide) Chemical Formula: $C_{36}H_{36}N_6O_2$ Molecular Weight: 584.72 Log P: 4.99	52
3.5	 *N*,*N*'-(heptane-1,7-diyl)bis(4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzamide) Chemical Formula: $C_{37}H_{38}N_6O_2$ Molecular Weight: 598.75 Log P: 5.41	83
3.6	 *N*,*N*'-(octane-1,8-diyl)bis(4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzamide) Chemical Formula: $C_{38}H_{40}N_6O_2$ Molecular Weight: 612.78 Log P: 5.83	74
3.7	 *N*,*N*'-(nonane-1,9-diyl)bis(4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzamide) Chemical Formula: $C_{39}H_{42}N_6O_2$ Molecular Weight: 626.81 Log P: 6.24	31
3.8	 *N*,*N*'-(decane-1,10-diyl)bis(4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzamide) Chemical Formula: $C_{40}H_{44}N_6O_2$ Molecular Weight: 640.83 Log P: 6.66	70

3.9	 *N*,*N*'-(1,4-phenylene)bis(4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzamide) Chemical Formula: $C_{36}H_{28}N_6O_2$ Molecular Weight: 576.66 Log P: 5.14	51
3.10	 N^4,$N^{4'}$-bis(4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)phenyl)-[1,1'-biphenyl]-4,4'-dicarboxamide Chemical Formula: $C_{42}H_{32}N_6O_2$ Molecular Weight: 652.76 Log P: 6.81	90

3A.3 Molecules in Library-2

3A.3.1 Design rationale of library-2 molecules

To investigate further the role of the single methylene spacer between the phenylene ring and the terminal benzimidazole ring (**Figure 3A.5**), library-2 molecules were designed without the phenylene ring and the single methylene spacer. For this purpose, three types of terminal building blocks were used to achieve the library-2 molecules, and these were further classified into three sub-groups, 2A, 2B and 2C with 2-methylquinoline, 6-fluoro-2-methylquinoline and 3*H*-1λ^4-indole terminal blocks, respectively.

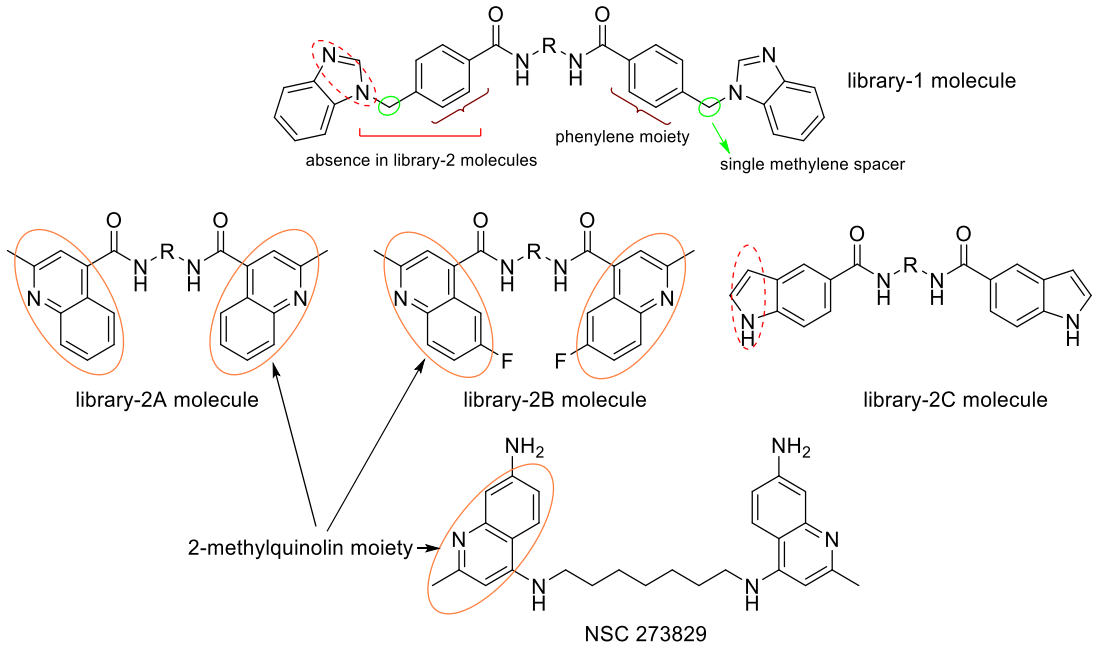


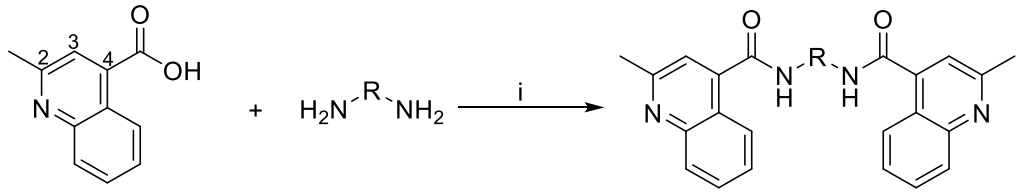
Figure 3A.7: Structural variation from library-1 molecules to library-2A, -2B and -2C molecules and features in common with NSC 273829.

To synthesize library-2A molecules, 2-methylquinoline was used as a terminal building block due to its structural similarity to the quinoline ring present in the NCI library compound NSC 273829. Similarly, 6-fluoro-2-methylquinoline was used as a terminal building block in library-2B molecules due to its structural similarity to the quinoline ring present in NSC 273829 and library-2A molecules. It has already been established that an intercalating unit should have a planar surface so that the molecule and/or the partial structure can sit in between the grooves like the filler of a sandwich. In this library of molecules, electronegative fluorine was used in the terminal block which would prevent the molecule from intercalating and thus provide a signal of importance of the single methylene spacer attached to the benzimidazole in library-1 molecules in kinking the structure in order to fit for intercalation. To pursue a similar investigation, 3*H*-1λ^4-indole, which is similar to the terminal building block in the library-1 molecules and commercially available, was used in the synthesis of the library-2C molecules.

3A.3.2 Synthesis of *bis*-quinoline type molecules in library-2A

Library-2A molecules were achieved by following a general HOBt/DIC-mediated amide coupling reaction in dimethylformamide (section **3A.1**). The reactions were very

straightforward and quick (6-7 h) compared with those of the library-1 molecules. Most of the compounds were purified by column chromatography using a polar solvent system (*e.g.*, dichloromethane/isopropanol, ethyl acetate/methanol) and achieved in a moderate yield (24-65%). A general synthetic scheme and list of compounds belonging to this library are given in **Figure 3A.8** and **Table 3A.2**, respectively.



3.11

Reaction conditions:
i) HOBt (0.83 mmol)/DIC (0.73 mmol) in DMF at RT, 6-7 h

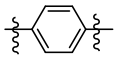
Here,
R = $-(CH_2)_n-$; 
n = 3, 6-10

Figure 3A.8: General synthetic scheme for library-2A molecules.

To achieve this library of molecules, 2-methylquinoline-4-carboxylic acid (0.44 mmol), HOBt (0.83 mmol) and DIC (0.73 mmol) were successively solubilized in dimethylformamide and allowed to stir at room temperature. After an initial activation step of approximately 30 minutes, the respective diamine (0.21 mmol) was added and allowed to react to completion to give the desired molecule. All the starting materials were soluble in dimethylformamide and reacted quickly, and thus most of the reaction had completed after 6-7 h. Distilled water was used to quench each reaction and the desired products were extracted with ethyl acetate (discussed in experimental section A.1.2). Most of the compounds were purified through flash column chromatography; for example, compounds **3.12** and **3.15** were obtained with dichloromethane/isopropyl alcohol (0-30%) and ethyl acetate/methanol (0-10%), respectively.

However, the purification of compound **3.14** was not so direct and was not purified after an initial attempt using n-hexane/ethyl acetate (0-80%). Afterwards, all the crude was collected

via flushing with ethyl acetate (100%). Finally, ethyl acetate/methanol (0-0.5%) was successfully used to elute the pure desired compound.

On the other hand, extraction and purification of compound **3.13** was long and difficult. During the extraction process, upon the addition of 15% lithium chloride in aqueous solution, the reaction mixture formed a fluffy, insoluble layer which was filtered and dissolved in methanol. As LC-MS showed some impurities in the methanol were contained in the product, and so petroleum ether was used to precipitate the product in methanol. Upon addition of a small amount of petroleum ether, the product precipitated in the round bottomed flask within an hour. Finally, the product was filtered, washed with water and freeze dried to achieve the pure compound.

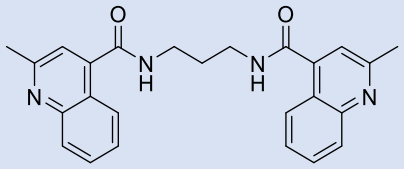
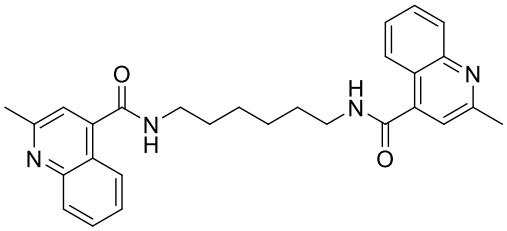
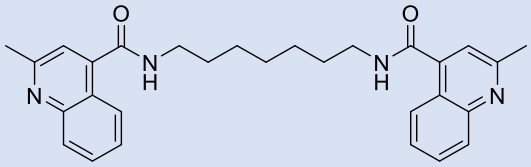
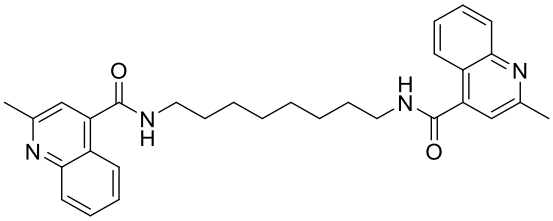
Compound **3.16** also formed a yellow precipitate during the extraction process with ethyl acetate; the precipitate was filtered and dissolved in methanol. Finally, the product was purified by flash column chromatography using ethyl acetate/methanol/2M ammonia solution in methanol (80/15/5).

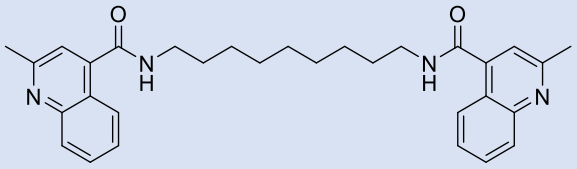
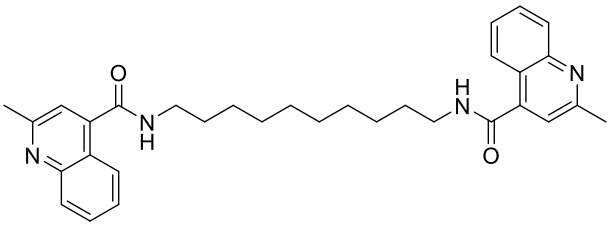
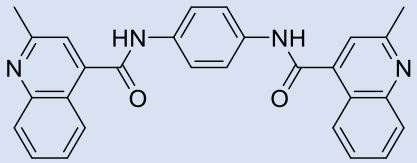
Extractions of compounds **3.17** and **3.18** were also tried with ethyl acetate, but the reaction mixture generated a fluffy layer during the extraction process. This was removed by filtration and freeze dried to give **3.17** and **3.18** in 38% and 55% yields, respectively.

However, compound **3.13** was obtained by recrystallization from ethyl acetate and 15% lithium chloride in aqueous solution, and compounds **3.17** and **3.18** were collected following the freeze drying of the ethyl acetate fraction of the crude reaction mixture.

Table 3A.2: *Bis*-quinoline molecules in library-2A.

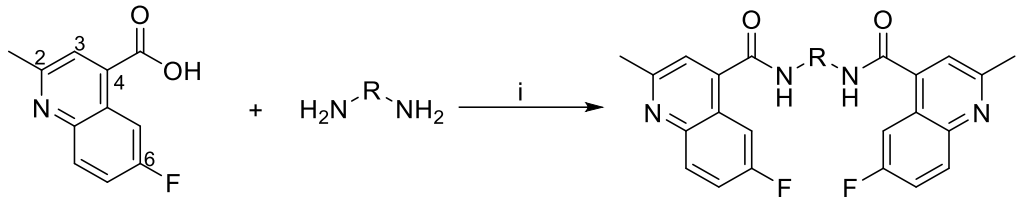
Compound	Structure (Chemical name, formula, MW, log*P*)	Yield (%)

3.12	 *N*,*N*'-(propane-1,3-diyl)bis(2-methylquinoline-4-carboxamide) Chemical Formula: $C_{25}H_{24}N_4O_2$ Molecular Weight: 412.49 Log P: 3.79	48
3.13	 *N*,*N*'-(hexane-1,6-diyl)bis(2-methylquinoline-4-carboxamide) Chemical Formula: $C_{28}H_{30}N_4O_2$ Molecular Weight: 454.57 Log P: 5.08	26
3.14	 *N*,*N*'-(heptane-1,7-diyl)bis(2-methylquinoline-4-carboxamide) Chemical Formula: $C_{29}H_{32}N_4O_2$ Molecular Weight: 468.60 Log P: 5.5	50
3.15	 *N*,*N*'-(octane-1,8-diyl)bis(2-methylquinoline-4-carboxamide) Chemical Formula: $C_{30}H_{34}N_4O_2$ Molecular Weight: 482.63 Log P: 5.91	24

3.16	 *N*,*N*'-(nonane-1,9-diyl)bis(2-methylquinoline-4-carboxamide) Chemical Formula: $C_{31}H_{36}N_4O_2$ Molecular Weight: 496.66 Log P: 6.33	64
3.17	 *N*,*N*'-(decane-1,10-diyl)bis(2-methylquinoline-4-carboxamide) Chemical Formula: $C_{32}H_{38}N_4O_2$ Molecular Weight: 510.68 Log P: 6.75	38
3.18	 *N*,*N*'-(1,4-phenylene)bis(2-methylquinoline-4-carboxamide) Chemical Formula: $C_{28}H_{22}N_4O_2$ Molecular Weight: 446.51 Log P: 5.22	55

3A.3.3 Synthesis of *bis*-(6-fluoro-2-methylquinoline) molecules in library-2B

Library-2 molecules were synthesized by using 6-fluoro-2-methylquinoline-4-carboxylic acid (0.24 mmol), HOBt (0.49 mmol) and DIC (0.43 mmol). As all of the starting materials were soluble in dimethylformamide (DMF), all reactions used to synthesize this library of molecules were carried out in DMF. Respective diamines (0.12 mmol) were added into the stirred reaction mixtures after the initial activation step. The general reaction scheme is shown below (**Figure 3A.9**). A total of six molecules were synthesized in this library, listed in **Table 3A.3.**



3.19

Reaction conditions:
i) HOBt (0.83 mmol)/DIC (0.73 mmol) in DMF at RT, 7.5-22.5 h

Here,
R = $-(CH_2)_n-$; n = 5-10

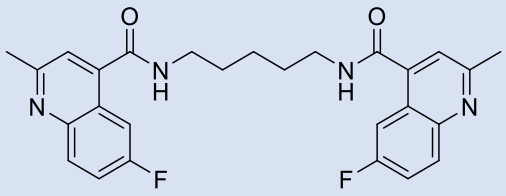
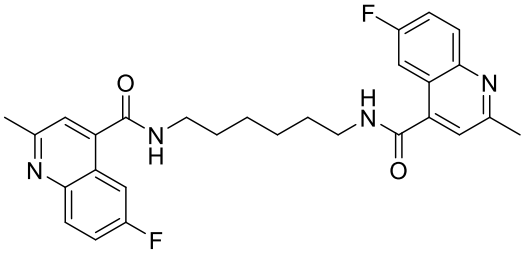
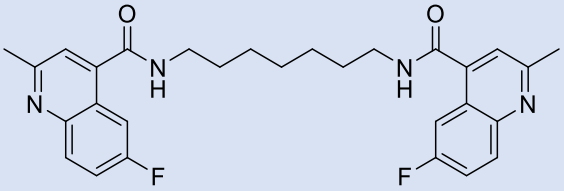
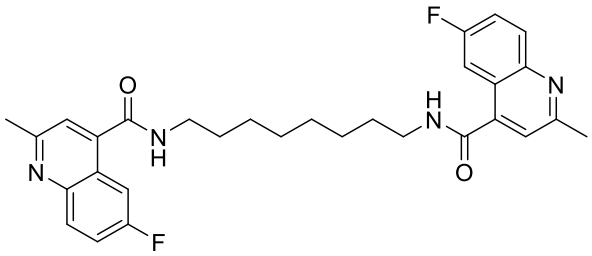
Figure 3A.9: General synthetic scheme for library-2B molecules.

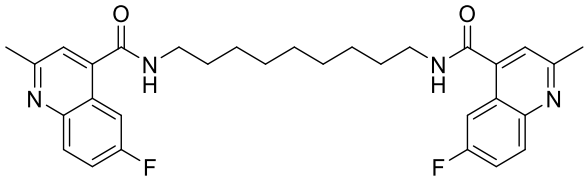
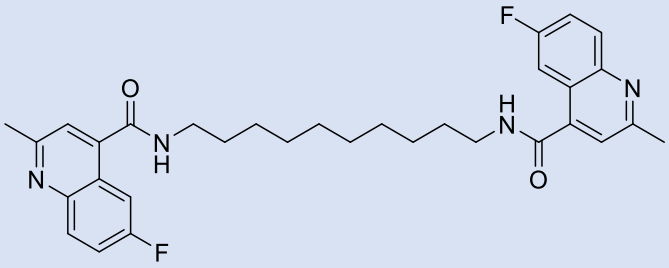
Water was used to quench each reaction and extractions were made with ethyl acetate. The combined organic phases were washed with a saturated aqueous solution of sodium bicarbonate, dried over magnesium sulphate, filtered and concentrated. Compound **3.23** was purified by flash column chromatography using n-hexane/ethyl acetate (0-70%).

The reactions for compounds **3.24** and **3.25** were longer, approximately 24 h. In both cases, additional acid (0.24 mmol) was added to force the reaction to completion, but there was no change in the rate of diamine consumption (by the acid). Finally, the reactions were finished after a day and products were purified by flash column chromatography. The reaction to synthesize compound **3.22** was finished after around 15 h; however, the initial attempt at purification with dichloromethane/ethyl acetate/methanol (97/2/1) was not successful. The desired product was eluted with impurities within first three column volumes of mobile phase. Then the combined organic phases were dried and purification was attempted by recrystallization with dichloromethane and diethyl ether. The process was successful though the yield was very low (25%).

Similarly, a recrystallization process was applied to purify compounds **3.20** and **3.21** by using dichloromethane and diethyl ether, with low yields achieved in both cases. The reaction times for compounds **3.20** and **3.21** were shorter (approximately 7.5 h) compared with the other members of this library, which was assumed to be a benign cause of poor yield and also indicated that not all of the starting materials were consumed in the reaction. As the required amounts of compounds were obtained, reactions for longer periods were not tried.

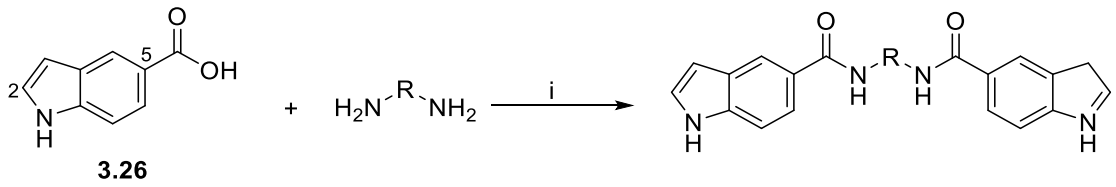
Table 3A.3: *Bis*-(6-fluoro-2-methylquinoline) molecules in library-2B.

Compound	Structure (Chemical name, formula, MW, log*P*)	Yield (%)
3.20	 *N*,*N*'-(pentane-1,5-diyl)bis(6-fluoro-2-methylquinoline-4-carboxamide) Chemical Formula: $C_{27}H_{26}F_2N_4O_2$ Molecular Weight: 476.53 Log P: 4.98	37
3.21	 *N*,*N*'-(hexane-1,6-diyl)bis(6-fluoro-2-methylquinoline-4-carboxamide) Chemical Formula: $C_{28}H_{28}F_2N_4O_2$ Molecular Weight: 490.55 Log P: 5.39	31
3.22	 *N*,*N*'-(heptane-1,7-diyl)bis(6-fluoro-2-methylquinoline-4-carboxamide) Chemical Formula: $C_{29}H_{30}F_2N_4O_2$ Molecular Weight: 504.58 Log P: 5.81	25
3.23	 *N*,*N*'-(octane-1,8-diyl)bis(6-fluoro-2-methylquinoline-4-carboxamide) Chemical Formula: $C_{30}H_{32}F_2N_4O_2$ Molecular Weight: 518.61 Log P: 6.23	92

Compound	Structure (Chemical name, formula, MW, log*P*)	Yield (%)
3.24	 *N*,*N*'-(nonane-1,9-diyl)bis(6-fluoro-2-methylquinoline-4-carboxamide) Chemical Formula: $C_{31}H_{34}F_2N_4O_2$ Molecular Weight: 532.64 Log P: 6.65	81
3.25	 *N*,*N*'-(decane-1,10-diyl)bis(6-fluoro-2-methylquinoline-4-carboxamide) Chemical Formula: $C_{32}H_{36}F_2N_4O_2$ Molecular Weight: 546.66 Log P: 7.06	90

3A.3.4 Synthesis of *bis*-indole molecules in library-2C

Library-2C molecules were synthesized to compare them with the library-1 molecules to find the importance of the phenylene in between the benzimidazole and amide functionalities of the library-1 molecules. In this library of molecules, the benzimidazole was replaced by an indole structure and the phenylene group was removed to establish the importance of the single methylene spacer in between the benzimidazole and phenylene. In this synthesis, 1*H*-indole-5-carboxylic acid was used as the side arms, which were connected by a methylene spacer through amide bonding. The general reaction scheme and list of compounds are presented in **Figure 3A.10** and **Table 3A.4**, respectively.



Reaction conditions:
i) HOBt (0.31 mmol)/DIC (0.27 mmol) in DMF at RT, 20-36 h

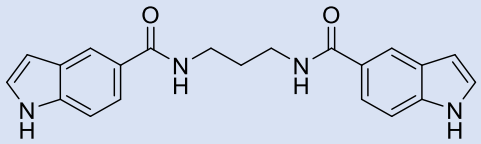
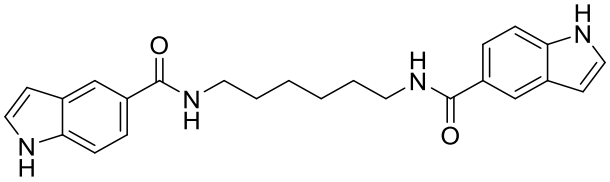
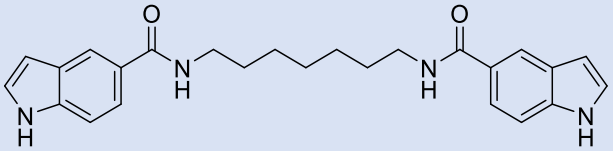
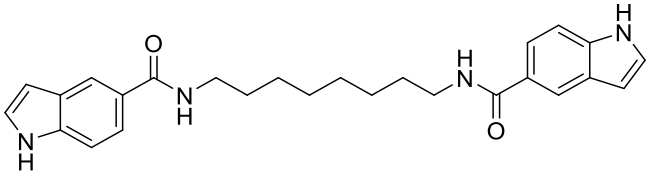
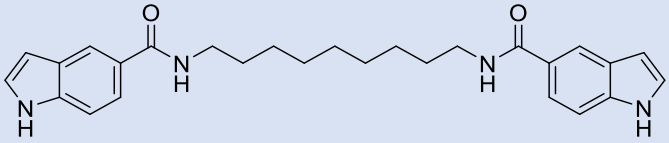
Here,
R = $-(CH_2)_n-$; n = 3, 6-9

Figure 3A.10: General synthetic scheme for library-2C molecules.

To synthesize the library-2C molecules, 1*H*-indole-5-carboxylic acid (0.16 mmol), HOBt (0.31 mmol) and DIC (0.27 mmol) were used and the reactions were carried out in dimethylformamide for better solubility. The respective diamines were added to the reaction mixture after the initial activation step. Similar to the previous library of molecules, reactions were quenched with water and extracted with ethyl acetate. The organic phases were combined, washed with a saturated aqueous solution of sodium bicarbonate, dried over magnesium sulphate, filtered and concentrated.

Reaction time was varied from reaction-to-reaction, but in most cases was 20-36 h. For example, compound **3.27** was formed after ~24 h reaction time, then extracted (discussed earlier) and purified by flash column chromatography with n-hexane/ethyl acetate (0-70%). The longest time was for synthesizing compound **3.31**, which was 36 h. All compounds were purified by flash column chromatography by using either n-hexane/ethyl acetate or methanol/ethyl acetate. Though the reaction times were longer than for the other molecules, the yields of most were good, being around 90% (except 73% for compound **3.29**).

Table 3A.4: *Bis*-indole molecules in library-2C.

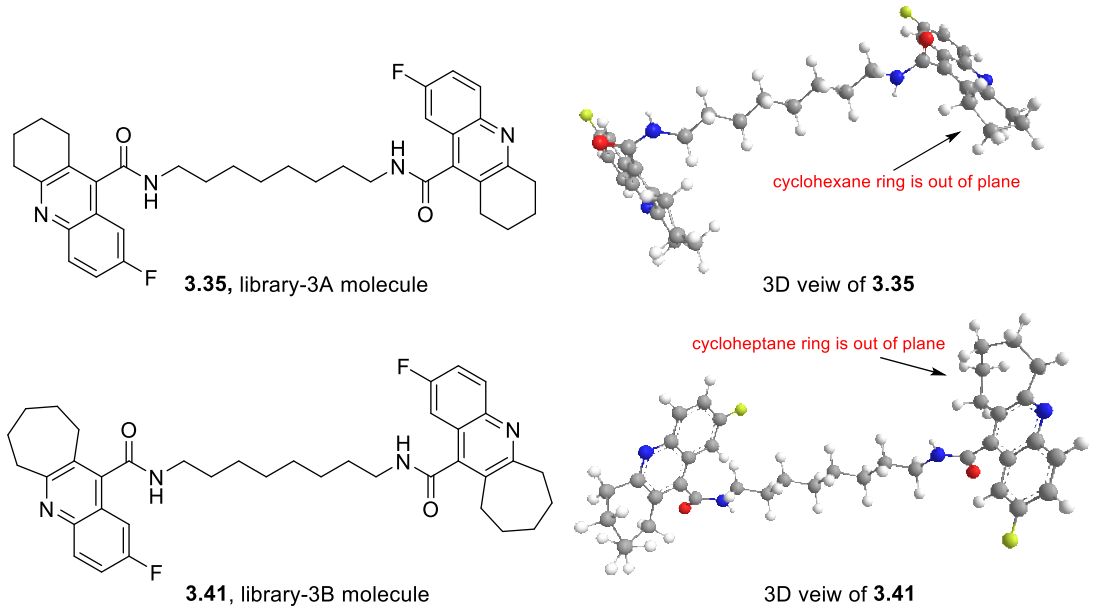
Compound	Structure (Chemical name, formula, MW, log*P*)	Yield (%)
3.27	 *N*,*N*'-(propane-1,3-diyl)bis(1*H*-indole-5-carboxamide) Chemical Formula: $C_{21}H_{20}N_4O_2$ Molecular Weight: 360.42 Log P: 1.3	93
3.28	 *N*,*N*'-(hexane-1,6-diyl)bis(1*H*-indole-5-carboxamide) Chemical Formula: $C_{24}H_{26}N_4O_2$ Molecular Weight: 402.50 Log P: 2.58	92
3.29	 *N*,*N*'-(heptane-1,7-diyl)bis(1*H*-indole-5-carboxamide) Chemical Formula: $C_{25}H_{28}N_4O_2$ Molecular Weight: 416.53 Log P: 3	73
3.30	 *N*,*N*'-(octane-1,8-diyl)bis(1*H*-indole-5-carboxamide) Chemical Formula: $C_{26}H_{30}N_4O_2$ Molecular Weight: 430.55 Log P: 3.42	94
3.31	 *N*,*N*'-(nonane-1,9-diyl)bis(1*H*-indole-5-carboxamide) Chemical Formula: $C_{27}H_{32}N_4O_2$ Molecular Weight: 444.58 Log P: 3.84	89

3A.4 Molecules in Library-3

3A.4.1 Design Rationale of Library-3 Molecules

For further investigation of structural features in the library-1 molecules, 7-fluoro-1,2,3,4-tetrahydroacridine (**3.32**) and 2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline moieties (**3.37**) were incorporated as the terminal blocks in library-3 molecules, where it was tried to keep structural correlation between the other library molecules and the lead NCI compound. The library-3 molecules were classified into two sub-groups, 3A and 3B, according to the terminal building blocks.

In the library-3A, terminal 7-fluoro-1,2,3,4-tetrahydroacridine moieties were used to generate some extent of bulkiness within the compound as well as to distort structural planarity due to the attached cyclohexane ring. This was hypothesized and confirmed through a 3D view of the molecule (**Figure 3A.11**) that the cyclohexane moiety would not be in the same plane as the methyl-quinoline moiety and thuswould make a kink in the ring structure and prevent the intercalation process. Similar approaches were used in the synthesis of library-3B molecules using 2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline moieties.



3.35, library-3A molecule

3D veiw of **3.35**

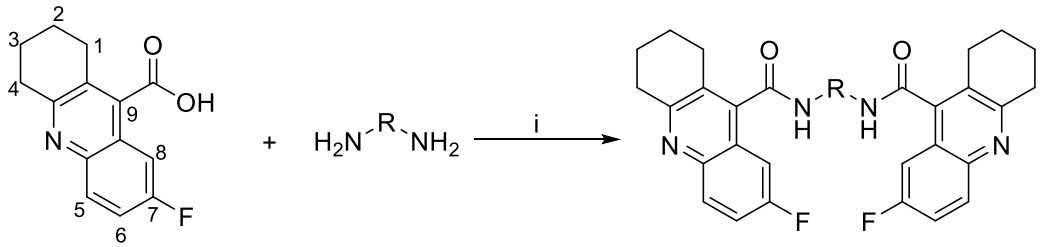
3.41, library-3B molecule

3D veiw of **3.41**

Figure 3A.11: Effect of the insertion of cyclohexyl and cycloheptyl rings in library-3A and -3B; 3D view shows the distortion in the molecule's planarity.

3A.4.2 Synthesis of *bis*-acridine type molecules in library-3A

In this library, a total of four compounds were synthesized using 7-fluoro-1,2,3,4-tetrahydroacridine-9-carboxylic acid (0.31 mmol), HOBt (0.61 mmol) and DIC (0.57 mmol).These starting materials were dissolved in dimethylformamide and allowed to stir to reach the activation stage. Afterwards, the respective diamines (0.15 mmol) were added and allowed to stir overnight (approximately 15 h). Reactions were straightforward and upon confirmation of the consumption of all the starting material, the reactions were quenched with water and subsequently extracted with ethyl acetate. Organic phases were combined and washed with an aqueous solution of sodium bicarbonate and dried over magnesium sulphate, filtered and concentrated. The general reaction scheme and compound details are reported in **Figure 3A.12** and **Table 3A.5**, respectively.



3.32

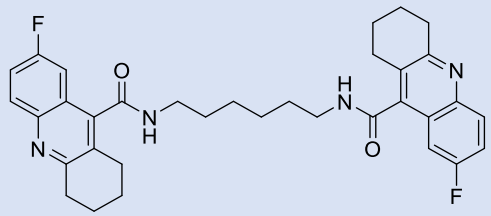
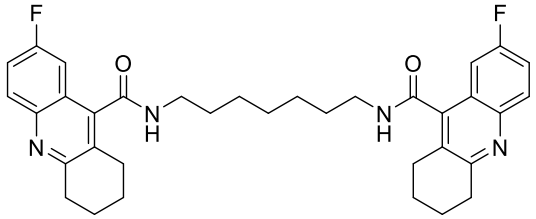
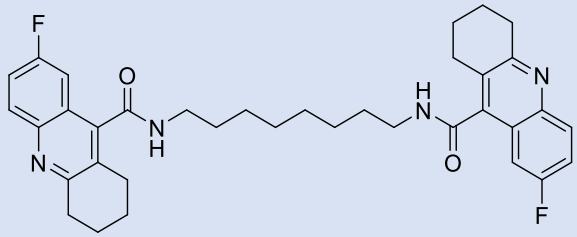
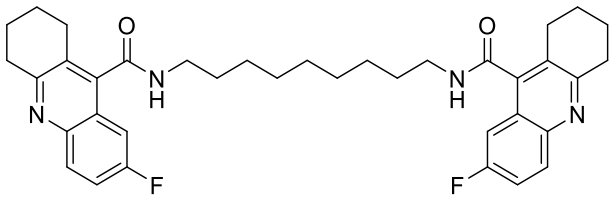
Reaction conditions:
i) HOBt (0.61 mmol)/DIC (0.57 mmol) in DMF at RT, 13.0-17.5 h

Here,
R = $-(CH_2)_n-$; n = 6-9

Figure 3A.12: General synthetic scheme for library-3A molecules.

All the compounds were purified by flash column chromatography with ethyl acetate/methanol (0-10%), but initially the purification was started with n-hexane/ethyl acetate (0-100%).

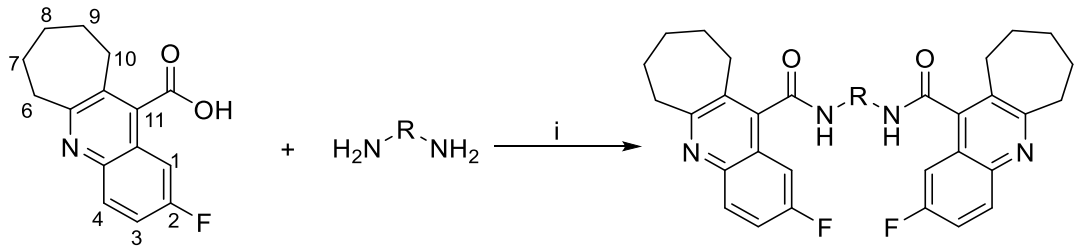
Table 3A.5: *Bis*-acridine type molecules in library-3A.

Compound	Structure (Chemical name, formula, MW, log*P*)	Yield (%)
3.33	 *N*,*N*'-(hexane-1,6-diyl)bis(7-fluoro-1,2,3,4-tetrahydroacridine-9-carboxamide) Chemical Formula: $C_{34}H_{36}F_2N_4O_2$ Molecular Weight: 570.68 Log P: 7.18	55
3.34	 *N*,*N*'-(heptane-1,7-diyl)bis(7-fluoro-1,2,3,4-tetrahydroacridine-9-carboxamide) Chemical Formula: $C_{35}H_{38}F_2N_4O_2$ Molecular Weight: 584.71 Log P: 7.6	39
3.35	 *N*,*N*'-(octane-1,8-diyl)bis(7-fluoro-1,2,3,4-tetrahydroacridine-9-carboxamide) Chemical Formula: $C_{36}H_{40}F_2N_4O_2$ Molecular Weight: 598.74 Log P: 8.01	57
3.36	 *N*,*N*'-(nonane-1,9-diyl)bis(7-fluoro-1,2,3,4-tetrahydroacridine-9-carboxamide) Chemical Formula: $C_{37}H_{42}F_2N_4O_2$ Molecular Weight: 612.77 Log P: 8.43	74

In the case of compound **3.34**, automated flash chromatography was used for purification with 4% methanol in ethyl acetate; this was assumed to be the cause of the low yield as higher percentages of methanol can elute the product faster, but with impurities.

3A.4.3 Synthesis of *bis*- (2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*] quinoline-11-carboxamide in library-3B

In this synthesis, 2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*] quinoline-11-carboxylic acid was used as terminal building blocks, connected by a methylene spacer through amide bonding. A suspension of 2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxylic acid (0.48 mmol) was stirred in dichloromethane with HOBt (0.96 mmol) and DIC (0.84 mmol) at room temperature. Dichloromethane was used as solvent for this class of molecule due to the acid's solubility in dichloromethane. After the initial activation step, the respective diamines (0.20 mmol) were added and the reactions left until completion. Upon confirmation of this by LC-MS and TLC, reactions were quenched with water and extracted with ethyl acetate. Organics were combined, washed with a saturated aqueous solution of sodium bicarbonate, dried over anhydrous magnesium sulphate, filtered and concentrated. A total of six compounds were synthesized in this library. The general reaction scheme and list of compounds are recorded in **Figure 3A.13** and **Table 3A.6**, respectively.



3.37

Reaction condition:
i) HOBt (0.96 mmol)/DIC (0.84 mmol) in DMF at RT, 17-36 h

Here,
R = $-(CH_2)_n-$; n = 5-10

Figure 3A.13: General synthetic scheme for library-3B molecules.

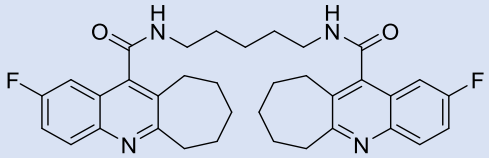
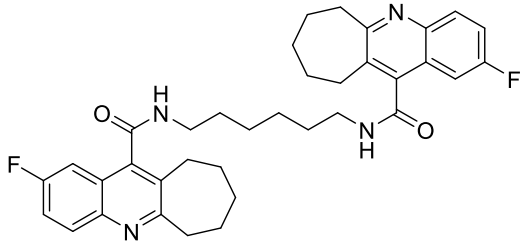
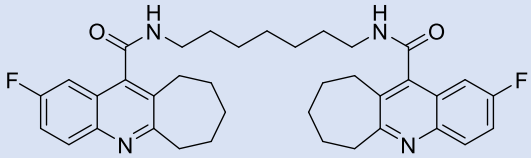
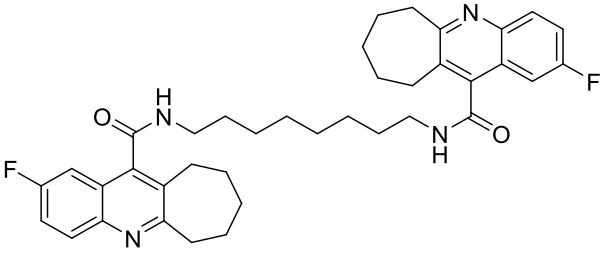
The first molecule of this library, **3.40**, could not be achieved as a *bis*-adduct, although the starting materials were consumed after 24 h reaction time. To synthesize the *bis*-amide-coupled material, an initial mono-amide product was purified by flash column chromatography using n-hexane/ethyl acetate (0-40%). Organics were combined and dried under vacuum. Thereafter, the mono-adduct compound was allowed to react again with additional acid (0.19 mmol), DIC (0.33 mmol) and HOBt (0.38 mmol). Overnight, the reaction reached completion and formation of the desired product was identified by LC-MS. Finally, purification was carried out by flash column chromatography with ethyl acetate/methanol (0-1%) to give **3.40** as a pale yellow solid. Similar difficulties were encountered during the synthesis of compound **3.39**.

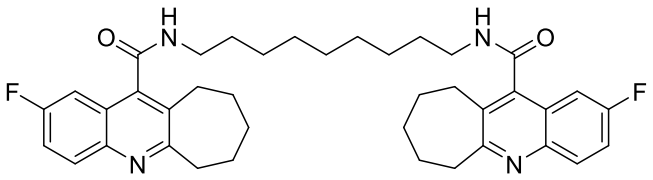
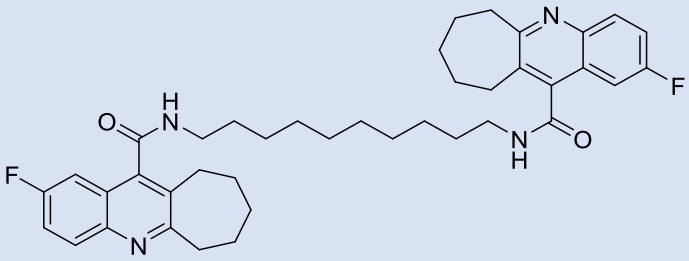
However, in this case the protocol followed for compound **3.40** did not work here; instead, a new reaction was carried out in dimethylformamide and, surprisingly, the reaction was finished after 17 h and the desired compound **3.39** was formed. The reaction was quenched with water and extracted with ethyl acetate. More importantly, the pure compound was obtained from the combined organic phases (ethyl acetate) without need of further purification.

The reaction to produce compound **3.38** took an exceptionally long time. In this synthesis, additional acid (0.48 mmol), HOBt (0.96 mmol) and DIC (0.84 mmol) were added to the reaction mixture after 36 h to force the reaction to completion. Finally, the reaction was completed after three days and the compound was purified by the 'Catch and release method' using a 10 g SCX-2 cartridge. However, none of the reactions required second time setup and/or addition of extra acid and coupling reagents, and most of them were completed after being left overnight (approximately 12-14 h).

Compounds **3.41**, **3.42** and **3.43** were obtained by following a similar purification technique as that followed for **3.40**.

Table 3A.6: *Bis*- (2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*] quinoline-11-carboxamide molecules in library-3B.

Compound	Structure (Chemical name, formula, MW, log*P*)	Yield (%)
3.38	 *N*,*N*'-(pentane-1,5-diyl)bis(2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxamide) Chemical Formula: $C_{35}H_{38}F_2N_4O_2$ Molecular Weight: 584.71 Log P: 7.6	84
3.39	 *N*,*N*'-(hexane-1,6-diyl)bis(2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxamide) Chemical Formula: $C_{36}H_{40}F_2N_4O_2$ Molecular Weight: 598.74 Log P: 8.01	73
3.40	 *N*,*N*'-(heptane-1,7-diyl)bis(2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxamide) Chemical Formula: $C_{37}H_{42}F_2N_4O_2$ Molecular Weight: 612.77 Log P: 8.43	94
3.41	 *N*,*N*'-(octane-1,8-diyl)bis(2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxamide) Chemical Formula: $C_{38}H_{44}F_2N_4O_2$ Molecular Weight: 626.79 Log P: 8.85	28

Compound	Structure (Chemical name, formula, MW, log*P*)	Yield (%)
3.42	 *N*,*N*'-(nonane-1,9-diyl)bis(2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxamide) Chemical Formula: $C_{39}H_{46}F_2N_4O_2$ Molecular Weight: 640.82 Log P: 9.27	70
3.43	 *N*,*N*'-(decane-1,10-diyl)bis(2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxamide) Chemical Formula: $C_{40}H_{48}F_2N_4O_2$ Molecular Weight: 654.85 Log P: 9.68	27

3A.5 Molecules in Library-4

3A.5.1 Design Rationale of Library-4 Molecules

To investigate whether the flexibility of the methylene spacers is essential to *bis*-intercalate with and stabilise the DNA/RNA hybrid duplex, and whether groove binding can be a possible mode of action to stabilise the DNA/RNA hybrid structure, library-4 molecules were synthesized using rigid and planar linkers. Additionally, six types of tertiary amine-containing side arms were utilized, assuming their conversion *in situ* into positively-charged species which would be attracted towards the minor groove of the DNA/RNA hybrid duplex sequence through electrostatic interactions.

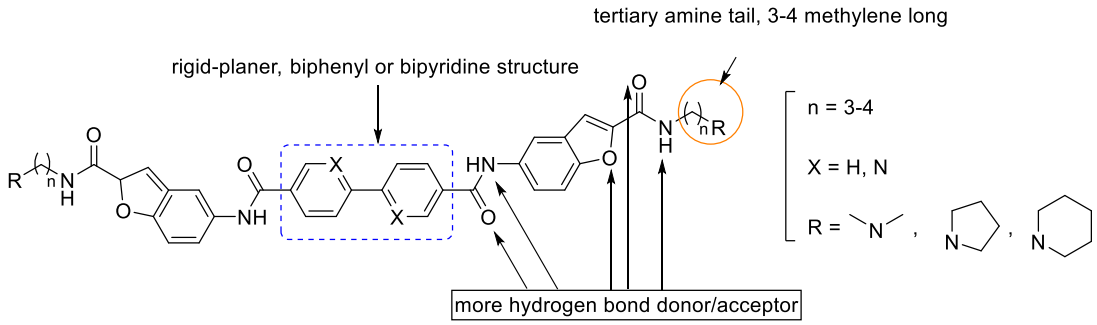


Figure 3A.14: Structural features of library-4 molecules.

3A.5.2 Synthesis of *bis*-benzofuran molecules in library-4

Library-4 molecules were synthesized to rationalize the importance of flexibility as a minor groove binding agent for the telomeric DNA/RNA hybrid structure. All the molecules in this library were synthesized using biphenylene and/or bipyridine spacers, where the carbon number within the spacer unit was the same as the 'hit' compound **3.6**, which was eight. A total of thirteen molecules were synthesized in this library (listed in **Table 3A.7**), by using [2,2′-bipyridine]-5,5′-dicarboxylic acid (**3.44**), [1,1′-biphenyl]-4,4′- dicarboxylic acid (**3.45**) and [2,2′-bipyridine]-4,4′-dicarboxylic acid (**3.46**) as spacer units to connect the benzofuran (**3.47**) structure at the 5-C position.

Additionally, six tertiary amine tails (**3.48-3.53, Figure 3A.15**) were also connected at the 2-C position of the benzofuran moiety (**3.47, Figure 3A.15**) by amide coupling reactions to achieve a monoamide-coupled product. To achieve the tertiary amine tail-linked product, three major types of tertiary amines were chosen (for example, dimethyl azane, pyrrolidine and pyridine) and the connecting methylene spacer was varied in length from 3-4 carbons.

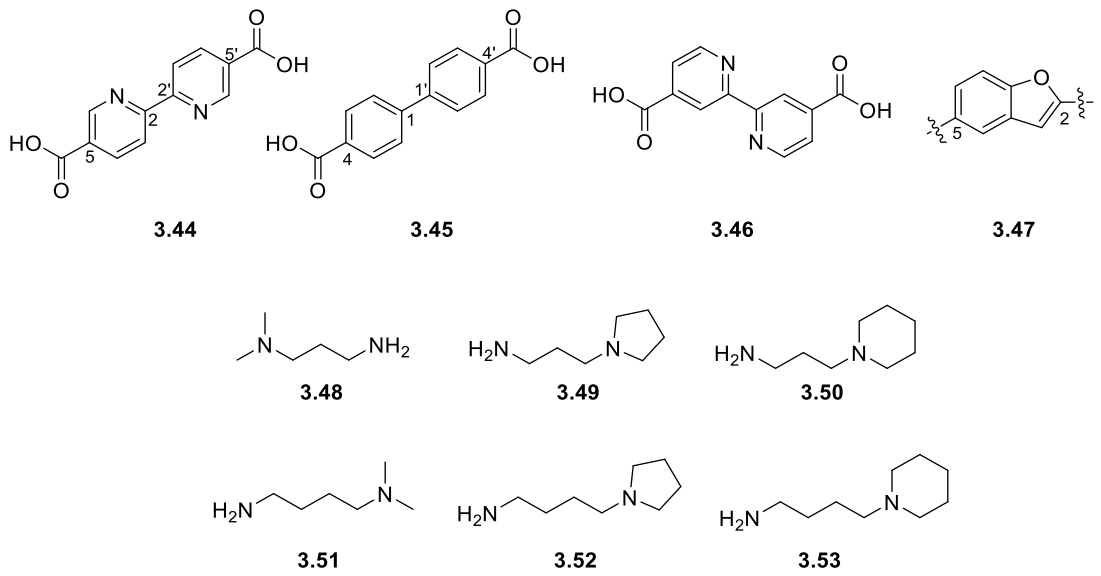


Figure 3A.15: Main building blocks used in the synthesis library-4 molecules.

The final compound was achieved after 3 steps of reaction. Initially, a solution of 5-nitrobenzofuran-2-carboxylic acid in dichloromethane was reacted with tertiary amine tails in the presence of HOBt and DIC. Every reaction finished after 1-2 h and the amide-coupled product was purified using SCX-2 cartridges by the 'Catch and release method' (see **section 2.3.2**).

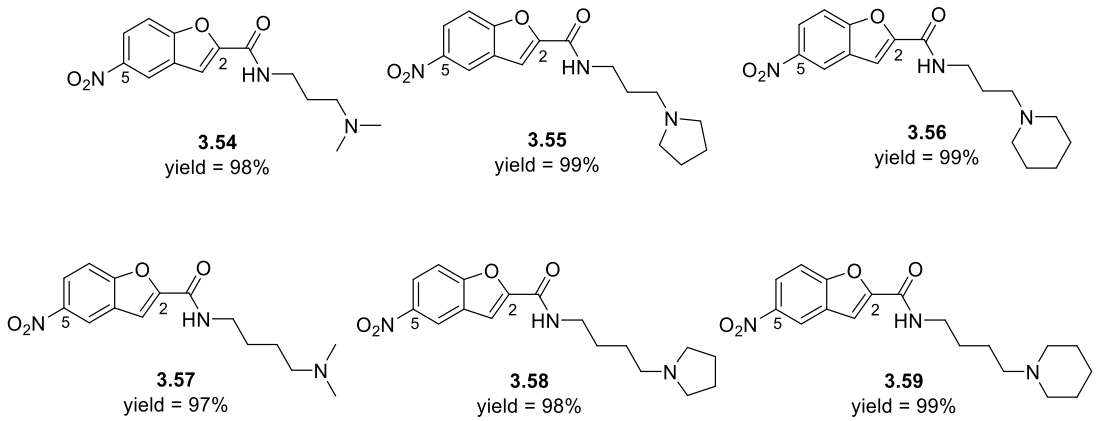


Figure 3A.16: Monoamide coupled product (1st step)

In the second step, a hydrogenation process was employed to obtain 5-NH_2 products of **3.54-3.59**. To achieve this, the reactant compound (**3.54-3.59, Figure 3A.16**) was dissolved in

either ethyl acetate or ethanol (maximum 150 mL) and a slurry of 10% Pd/C (10% w/w compared with the reactant) in the same solvent was added. The mixture was hydrogenated in a Parr Hydrogenator at 35-40 mbar until the H_2 uptake was finished. The progress of the reaction was monitored by either TLC or LC-MS. The reaction mixture was filtered through a layer of celite (diatomaceous earth) and the residue washed with either ethyl acetate or ethanol. The solvent was removed under vacuum using a rotary evaporator to obtain the reduced product (**3.60-3.65, Figure 3A.17**).

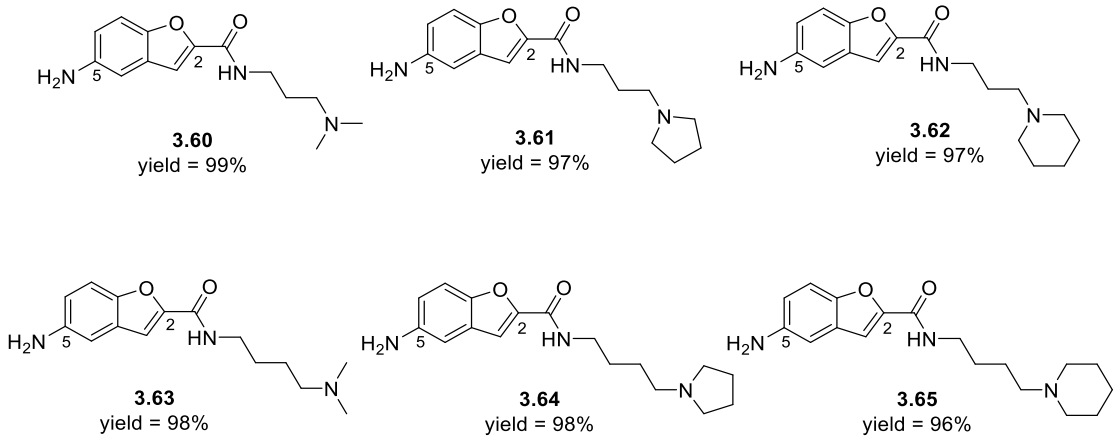
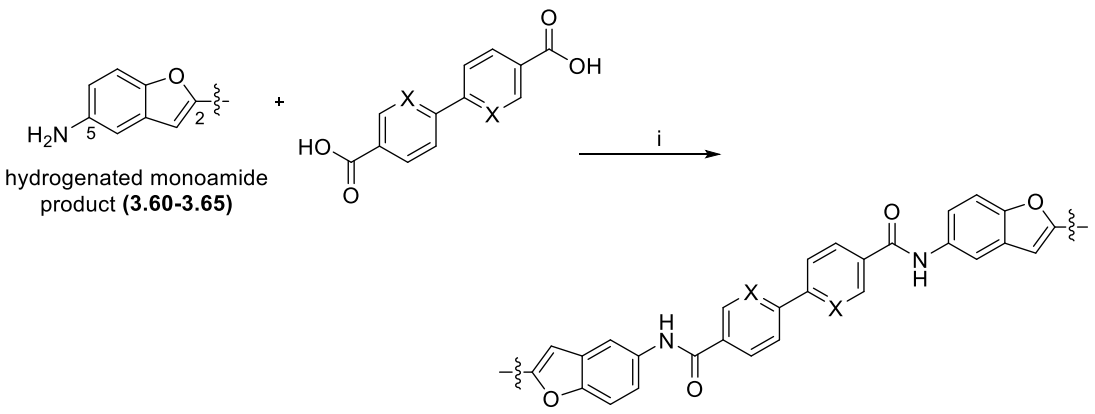


Figure 3A.17: Hydrogenated monoamide-coupled product (2nd step)

In the final step, the hydrogenated monoamide-coupled product was reacted with either [2,2′-bipyridine]-5,5′-dicarboxylic acid (**3.44**) or [1,1′-biphenyl]-4,4′-dicarboxylic acid (**3.45**) to achieve the *bis*-amide molecules. Only one compound, **3.68**, was synthesized by using [2,2′-bipyridine]-4,4′-dicarboxylic acid (**3.46**) as a spacer to evaluate the correlation between the curvature of the compound and binding affinity towards the telomeric DNA/RNA hybrid duplex sequence. The general reaction scheme is described in **Figure 3A.18**.



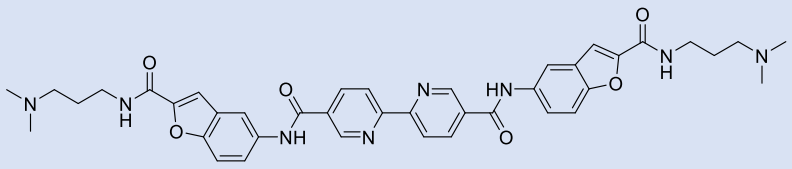
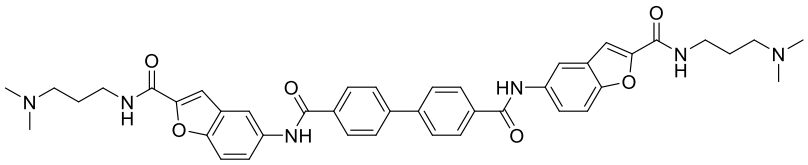
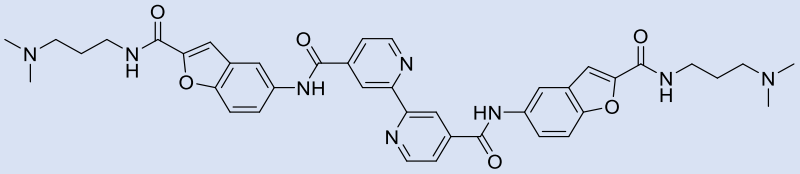
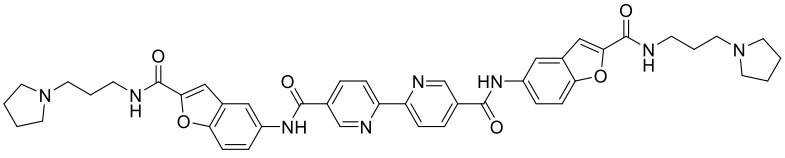
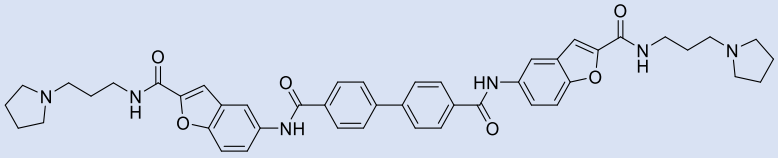
Reaction condition:
i) HOBt (0.49 mmol)/DIC (0.43 mmol) in DMF at RT, 13-15 h

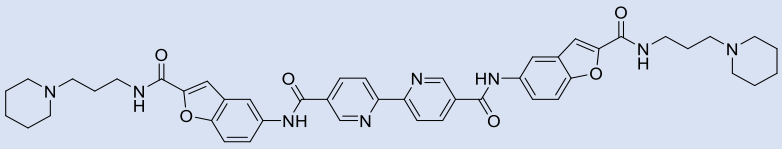
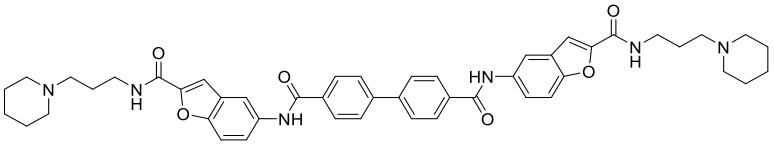
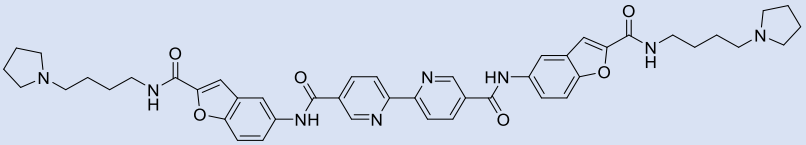
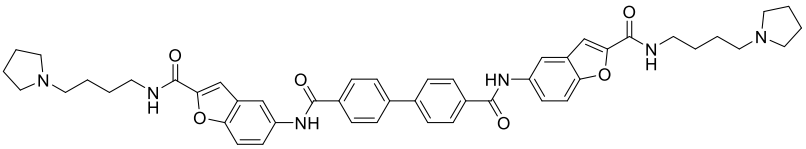
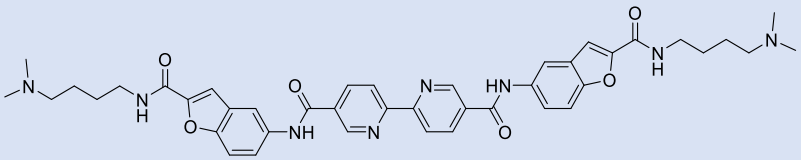
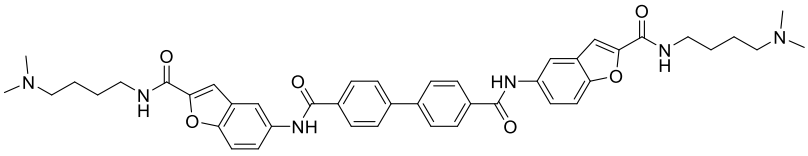
Here, X = N, H

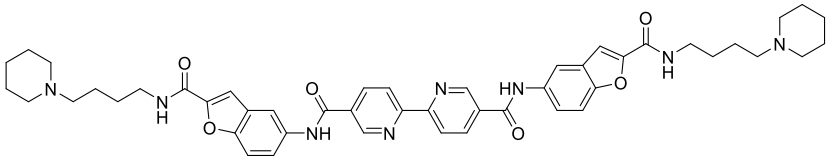
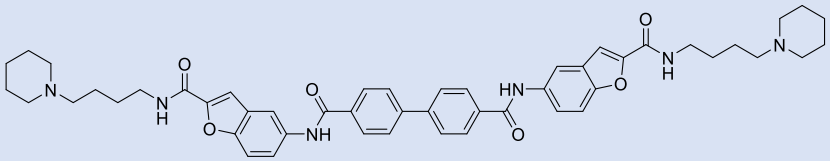
Figure 3A.18: General scheme of reaction involved in the synthesis of library-4 molecules.

Library-4 molecules were difficult to purify though all of them have a tertiary amine, which is believed to be stacked in the strong cation exchange sorbent within SCX-2 cartridges. In an initial attempt, the reaction mixture was poured through the activated cartridge and impurities allowed to elute out. However, it was observed that the cartridge became blocked by the stacked molecules within it, and thus did not allow any solvent to pass through. In the preliminary trial, the cartridge needed to be broken to obtain the slurry of silica mixed with the desired compound, which was then re-dissolved in methanol and filtered to collect the product. However, the product was not pure enough as the methanol dissolved a small amount of silica, which was mixed with the compound. To avoid this problem, compounds were purified in a different way. Firstly, the reaction mixture was placed in a round bottomed flask and a small amount of toluene (5-6 mL) was added to it. Subsequently, n-hexane was added in excess amount to form a cloudy solution that finally created a white precipitate of compound. Afterwards, the precipitate was collected on a sintered funnel and roughly dried under vacuum. LC-MS confirmed the respective product mass-to-charge (m/z) ratio, which ultimately confirmed the success of the purification. To achieve the highest compound purity, the product was washed with acetonitrile/water (50/50) and freeze dried. LC-MS, ^{1}H NMR, ^{13}C NMR, IR and HRMS analyses were carried out to check the compound's quality and purity.

Table 3A.7: *Bis*-benzofuran type molecules in library-4.

Compound	Structure (Chemical name, formula, MW, log*P*)	Yield (%)
3.66	 $N^5,N^{5'}$-bis(2-((3-(dimethylamino)propyl)carbamoyl)benzofuran-5-yl)-[2,2'-bipyridine]-5,5'-dicarboxamide Chemical Formula: $C_{40}H_{42}N_8O_6$ Molecular Weight: 730.83	24
3.67	 $N^4,N^{4'}$-bis(2-((3-(dimethylamino)propyl)carbamoyl)benzofuran-5-yl)-[1,1'-biphenyl]-4,4'-dicarboxamide Chemical Formula: $C_{42}H_{44}N_6O_6$ Molecular Weight: 728.85	50
3.68	 $N^4,N^{4'}$-bis(2-((3-(dimethylamino)propyl)carbamoyl)benzofuran-5-yl)-[2,2'-bipyridine]-4,4'-dicarboxamide Chemical Formula: $C_{40}H_{42}N_8O_6$ Molecular Weight: 730.83	26
3.69	 $N^5,N^{5'}$-bis(2-((3-(pyrrolidin-1-yl)propyl)carbamoyl)benzofuran-5-yl)-[2,2'-bipyridine]-5,5'-dicarboxamide Chemical Formula: $C_{44}H_{46}N_8O_6$ Molecular Weight: 782.90	64
3.70	 $N^4,N^{4'}$-bis(2-((3-(pyrrolidin-1-yl)propyl)carbamoyl)benzofuran-5-yl)-[1,1'-biphenyl]-4,4'-dicarboxamide Chemical Formula: $C_{46}H_{48}N_6O_6$ Molecular Weight: 780.93	38

Compound	Structure (Chemical name, formula, MW, log*P*)	Yield (%)
3.71	 $N^5,N^{5'}$-bis(2-((3-(piperidin-1-yl)propyl)carbamoyl)benzofuran-5-yl)-[2,2'-bipyridine]-5,5'-dicarboxamide Chemical Formula: $C_{46}H_{50}N_8O_6$ Molecular Weight: 810.96	48
3.72	 $N^4,N^{4'}$-bis(2-((3-(piperidin-1-yl)propyl)carbamoyl)benzofuran-5-yl)-[1,1'-biphenyl]-4,4'-dicarboxamide Chemical Formula: $C_{48}H_{52}N_8O_6$ Molecular Weight: 808.98	55
3.73	 $N^5,N^{5'}$-bis(2-((4-(pyrrolidin-1-yl)butyl)carbamoyl)benzofuran-5-yl)-[2,2'-bipyridine]-5,5'-dicarboxamide Chemical Formula: $C_{46}H_{50}N_8O_6$ Molecular Weight: 810.96	51
3.74	 $N^4,N^{4'}$-bis(2-((4-(pyrrolidin-1-yl)butyl)carbamoyl)benzofuran-5-yl)-[1,1'-biphenyl]-4,4'-dicarboxamide Chemical Formula: $C_{48}H_{52}N_8O_6$ Molecular Weight: 808.98	48
3.75	 $N^5,N^{5'}$-bis(2-((4-(dimethylamino)butyl)carbamoyl)benzofuran-5-yl)-[2,2'-bipyridine]-5,5'-dicarboxamide Chemical Formula: $C_{42}H_{46}N_8O_6$ Molecular Weight: 758.88	62
3.76	 $N^4,N^{4'}$-bis(2-((4-(dimethylamino)butyl)carbamoyl)benzofuran-5-yl)-[1,1'-biphenyl]-4,4'-dicarboxamide Chemical Formula: $C_{44}H_{48}N_8O_6$ Molecular Weight: 756.90	58

Compound	Structure (Chemical name, formula, MW, log*P*)	Yield (%)
3.77	 $N^5,N^{5'}$-bis(2-((4-(piperidin-1-yl)butyl)carbamoyl)benzofuran-5-yl)-[2,2'-bipyridine]-5,5'-dicarboxamide Chemical Formula: $C_{48}H_{54}N_8O_6$ Molecular Weight: 839.01	57
3.78	 $N^4,N^{4'}$-bis(2-((4-(piperidin-1-yl)butyl)carbamoyl)benzofuran-5-yl)-[1,1'-biphenyl]-4,4'-dicarboxamide Chemical Formula: $C_{50}H_{56}N_6O_6$ Molecular Weight: 837.03	49

Chapter 3B: Biophysical and Biological Evaluation: Results and Discussion

3B.1 Fluorescence Resonance Energy Transfer (FRET) Assay

Firstly, the library-1 molecules were assessed against the telomeric DNA/RNA hybrid duplex (DRH) sequence (TTA-GGG-TTA-GGG-TTT-TTT-CCC-UAA-CCC-UAA) along with a control cDRH DNA sequence (a hairpin DNA sequence in which the uridine (**U**) of the DRH duplex sequence had been replaced with thymidine (**T**)). According to the preliminary biophysical observations (see **Table-3B.1**), compound **3.6** was the most selective towards the DRH duplex compared with the cDRH (where RNA bases of DRH were replaced by corresponding DNA bases) structure (*i.e.*, a 9-folds selectivity), stabilizing the heteroduplex by ΔT_m values of 9.5 and 7.2 °C at 2 μM and 1 μM ligand concentrations, respectively. In the case of library-1 molecules, chains of 3 and 5-10 methylene units were used for the central spacer, which connected to the 1-benzyl-1*H*-benzo[*d*]imidazole units via amide linkages (**Figure 3B.1**).

Compared with NSC 273829, compound **3.6** was modified by changing the length of the spacer, in addition to including phenyl (compound **3.9**) and biphenyl (compound **3.10**) units to determine the impact of flexibility and rigidity on binding to the minor groove of telomeric DNA/RNA hybrid duplex sequences. Interestingly, it had been known for some time that there was a correlation between the linker length connecting the 1-benzyl-1*H*-benzo[*d*]imidazole units and DRH stabilizing activity, and this was further confirmed by molecular modelling and dynamics studies (**section 3B.2**).

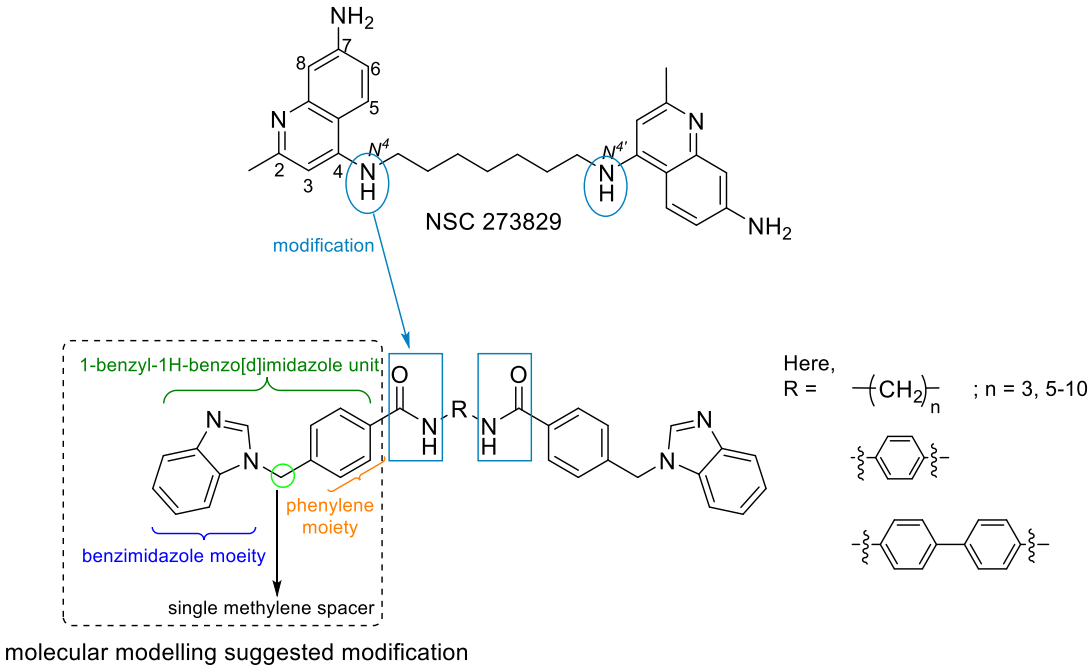
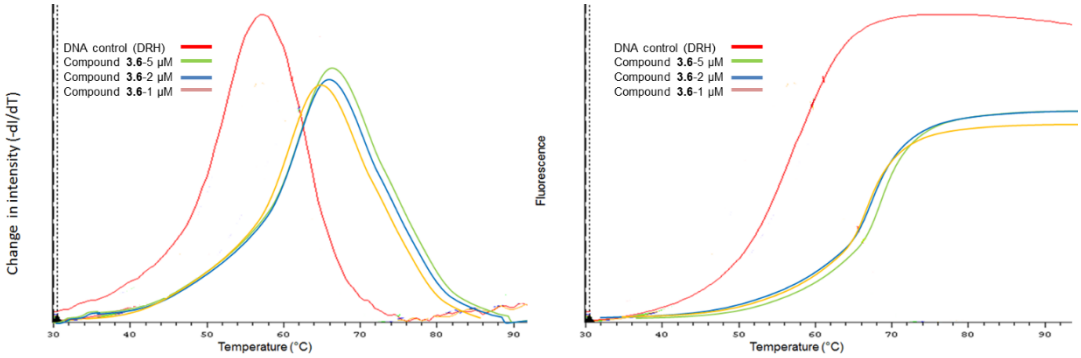


Figure 3B.1: Structural features and highlighted modifications in library-1 molecules.

According to **Table 3B.1** (*i.e.*, data from the FRET melting assay), DRH stabilizing activity could be ranked in the order **3.6** > **3.8** > **3.5** > **3.7** > **3.10** > **3.3** > **3.4** > **3.2** > **3.9** at a 1 μM ligand concentration, whereas stabilizing activity for cDRH was very poor. In a DNA targeting drug design study, selectivity of novel molecules towards the targeted DNA compared with control duplex DNA is desirable to eliminate unwanted toxicity.

In this study, a duplex hairpin DNA structure (labelled as cDRH) was used as a control, and to evaluate selectivity toward the targeted DNA/RNA hybrid duplex. It was observed that **3.4**, **3.6 (Figure 3B.2)** and **3.8** were the most selective toward the DNA/RNA hybrid duplex compared with the control sequence. In molecules **3.9** and **3.10**, the flexible methylene spacer had been replaced with the planar phenyl (4-carbon distance) and biphenyl units (8-carbon distance) to establish the impact of flexibility or rigidity of structure on DNA-drug binding. However, **3.9** and **3.10** did not show either improved selectivity or stabilization towards the DNA/RNA hybrid duplex. Molecular modelling and molecular dynamics studies suggested the possible reasons for the lower activities of **3.9** and **3.10** (see Section **3B.2.1**).



(a) (b)

Figure 3B.2: FRET melting curve for compound **3.6** at three concentrations (5 μM, 2 μM and 1 μM), where (a) change of intensity of fluorescence over the change of temperature (-dI/dT) and (b) increment of fluorescence due to increase of temperature.

Table 3B.1: FRET melting assay results for library-1 molecules with telomeric DNA/RNA hybrid duplex (DRH) and control cDRH sequences.

Compound	ΔT_m (°C)				Selectivity (DRH/cDRH)	
	DRH		cDRH			
	2 μM	1 μM	2 μM	1 μM	2 μM	1 μM
3.2	2.4±0.2	0.6±0.0	1.8±0.1	0.5±0.0	1.3	1.2
3.3	5.4±0.2	2.4±0.0	1.5±0.0	0.6±0.0	3.6	4.0
3.4	5.6±0.3	2.3±0.0	3.3±0.1	0.2±0.0	1.7	11.5
3.5	8.6±0.2	4.5±0.0	3.2±0.2	1.3±0.0	2.7	3.5
3.6	9.5±0.1	7.2±0.0	1.3±0.1	0.8±0.0	7.3	9.0
3.7	6.8±0.3	4.5±0.5	2.3±0.2	1.2±0.0	2.9	3.7
3.8	8.2±0.1	6.8±0.0	1.4±0.3	0.9±0.0	5.9	7.6
3.9	1.2±0.0	0.5±0.0	0.8±0.2	0.3±0.0	1.5	1.7
3.10	3.6±0.1	3.2±0.0	1.2±0.4	0.6±0.0	3.0	5.3

Library-1 molecules were also assessed against the telomeric G-quadruplex sequence (hTelo) F21T. A dual-labelled oligonucleotide sequence that forms an intramolecular G-quadruplex structure in the presence of a monovalent cation (Na^+/K^+) was used for the quadruplex

stabilization assay. In this study, the FAM-TAMRA dual-labelled oligomer (5´d-GGG-TTA-GGG-TTA-GGG-TTA-GGG-3´) was used to represent the F21T DNA sequence. Library-1 molecules provided a good stabilizing efficacy for the F21T sequence, with ΔT_m values ranging from 5.8-9.0 °C and 4.5-6.5 °C at 2 μM and 1 μM ligand concentrations, respectively (**Table 3B.2**). Compound **3.6** showed moderate stabilization of the hTelo sequence (ΔT_m = 8.8 and 6.5 °C at 2 μM and 1 μM ligand concentrations, respectively), which suggested that it might be a good candidate for further development as a dual mechanism telomerase inhibitor that could stabilize both the telomeric DNA/RNA hybrid duplex and the hTelo sequence (either by intercalating and/or minor groove binding).

Table 3B.2: FRET melting assay results for library-1 molecules with telomeric G-quadruplex (F21T) and control duplex DNA (cDDS) sequences.

Compound	ΔT_m (°C)				Selectivity (F21T/cDDS)	
	F21T		cDDS			
	2 μM	1 μM	2 μM	1 μM	2 μM	1 μM
3.2	3.5±0.0	1.4±0.0	1.0±0.1	0.8±0.0	3.5	1.7
3.3	8.8±0.1	5.8±0.0	1.3±0.2	1.0±0.0	6.8	5.8
3.4	5.8±0.3	4.5±0.0	1.2±0.0	0.8±0.0	4.8	5.6
3.5	8.5±0.0	5.5±0.0	1.2±0.2	1.0±0.0	7.1	5.5
3.6	8.8±0.4	6.5±0.0	1.0±0.0	0.7±0.0	8.8	9.3
3.7	9.0±0.2	6.5±0.0	1.5±0.2	0.9±0.0	6.0	7.2
3.8	7.8±0.2	6.2±0.0	1.5±0.2	0.8±0.0	6.0	7.5
3.9	1.6±0.2	0.2±0.0	1.1±0.1	0.3±0.0	1.5	0.7
3.10	1.5±0.2	0.5±0.0	1.0±0.3	0.5±0.0	1.5	1.0

However, upon changing the methylene spacer to a planar structure with non-flexible phenylene and biphenylene spacers (*i.e.*, **3.9** and **3.10**), the stabilization properties dropped to 0.2 and 0.5 °C, respectively, at the same concentration. Interestingly, reducing the

methylene spacer to a three-carbon length (**3.2**) also caused a noticeable decrease in stabilizing effect (*i.e.*, ΔT_m = 1.4 °C at 1 μM concentration).

The other library molecules were also evaluated by FRET melting assay against telomeric DNA/RNA hybrid duplexes (DRH), telomeric G-quadruplex sequences (F21T) and two control DNA duplex sequences (*i.e.*, cDRH and cDDS) at 5, 2 and 1 μM ligand concentrations. However, none provided stabilization of ≥0.5 °C at the highest ligand concentration (*i.e.*, 5 μM) toward the telomeric DNA/RNA hybrid duplex.

Interestingly, the molecules in library-3B showed moderate DRH stabilization properties (*i.e.*, ΔT_m = 3.2-6.5 °C for 2 μM conc. and ΔT_m = 2.3-3.5 °C for 1 μM conc.), while a 2 μM concentration of library-3B molecules also stabilized the cDRH sequence by a significant amount (*i.e.*, ΔT_m = 1.3-3.3 °C). Thus, the FRET melting result for library-3B molecules indicated that they were non-selective binders of the DRH sequence. Individual FRET melting temperatures are listed in **Table 3B.3.**

Table 3B.3: FRET melting assay results for library-3B molecules with telomeric DNA/RNA hybrid duplex (DRH) and control cDRH sequences.

Compound	ΔT_m (°C)				Selectivity (DRH/cDRH)	
	DRH		cDRH			
	2 μM	1 μM	2 μM	1 μM	2 μM	1 μM
3.38	6.5±0.1	3.5±0.0	2.3±0.0	0.5±0.0	2.8	7.0
3.39	5.5±0.2	3.3±0.0	3.3±0.0	0.8±0.0	1.7	4.1
3.40	3.2±0.0	2.3±0.0	2.3±0.1	1.2±0.0	1.4	1.9
3.41	5.5±0.3	3.4±0.0	3.2±0.0	1.3±0.0	1.7	2.6
3.42	6.5±0.1	3.2±0.0	1.3±0.1	0.8±0.0	5.0	4.0
3.43	4.7±0.2	2.3±0.0	2.2±0.1	1.2±0.0	2.1	1.9

In addition, the library-3B molecules caused moderate stabilization of the hTelo DNA sequence (F21T) with ΔT_m values of 3.5-7.5 °C (at a 2 μM ligand concentration) and 2.0-4.5 °C (at a 1 μM ligand concentration), which are presented in **Table 3B.4**. According to the FRET

melting data for library-3B molecules, it was observed that **3.38** (*i.e.*, contains a five methylene spacer) was selective towards both DRH and hTelo sequences compared with the respective control DNA sequences at a 1 μM ligand concentration, which suggested further investigation.

Table 3B.4: FRET melting assay results for library-3B molecules with telomeric G-quadruplex (F21T) and control duplex DNA (cDDS) sequences.

Compound	ΔT_m (°C)				Selectivity (F21T/cDDS)	
	F21T		cDDS			
	2 μM	1 μM	2 μM	1 μM	2 μM	1 μM
3.38	4.8±0.2	4.5±0.0	1.2±0.2	0.7±0.0	4.0	6.4
3.39	6.5±0.1	4.5±0.0	1.4±0.2	0.9±0.0	4.6	5.0
3.40	3.5±0.0	2.5±0.0	1.1±0.2	0.6±0.0	3.2	4.2
3.41	7.5±0.6	2.7±0.5	1.0±0.0	0.5±0.0	7.5	5.4
3.42	6.5±0.2	3.5±0.0	1.5±0.2	0.7±0.0	4.3	5.0
3.43	5.3±0.3	2.0±0.0	1.3±0.0	0.8±0.0	4.1	2.5

A double-labelled 45-base pair DNA sequence was used in this study as a control (cDDS, Section **5B.1.1**). As the designed molecules of library-1 were synthesized based on the NCI diversity set, it was expected that the synthesized analogues would not stabilize duplex DNA. However, a few of them did, at the lower ΔT_m values of 0.5-1.5 °C for a 1 μM ligand concentration. It was assumed that the dimethyl sulfoxide (DMSO) used to dissolve the molecules prior to assay may cause some degree of stabilization, so control experiments were carried out. It was also confirmed by FRET melting experiments that the lead compounds **3.3**, **3.6** and **3.8** gave ΔT_m values of 0.2-0.9 °C at a 1 μM concentration with control duplex DNA indicating a profound selectivity for DRH sequences versus duplex DNA sequences.

To study the impact of the spacer unit on affinity for the telomeric DNA/RNA hybrid duplex sequence, several new molecules were synthesized. According to FRET results, ΔT_m values were reduced by approximately 50% when the methylene spacer was increased or decreased in length from the optimum. Compound **3.6** was also transformed into a rigid structure by

replacing the flexible methylene chain with π-electron-containing phenylene and biphenylene moieties (*e.g.*, **3.9** and **3.10** respectively). For these molecules, only a 0.2-0.5 °C increase in ΔT_m was observed for the telomeric DNA/RNA hybrid duplex with a much higher affinity for duplex DNA. The conformation and geometry of standard duplex DNA is much wider than that of A-form hybrid duplex DNA, in which the minor grooves are much narrower and tighter, which could explain the observed results. Additionally, in order to intercalate between nucleobases, a molecule needs to be either crescent-shaped or more planar with the presence of a π-system. Thus the results suggest that the 8-methylene spacer is the ideal length for binding to the targeted hybrid duplex DNA sequence and allows the molecule to *bis*-intercalate into the nucleobases of the hybrid duplex. Moreover, this linker length is sufficiently flexible towards the minor groove of the hybrid duplex and allows the molecule to wrap around the hybrid structure as well as intercalate in-between nucleobases within the terminal planar π-system scaffold.

3B.2 Molecular Modelling and Molecular Dynamics (MD) Studies[2]

Molecular modelling comprises both theoretical methods and computational techniques that can be employed together to model or imitate the behaviour of molecules toward specific DNA sequences or protein structures. They are used in the areas of drug design and computational biology for studying molecular systems, where molecules ranging from small chemical scaffolds to large biological macromolecules can be analysed. A common feature of molecular modelling techniques is the atomistic level of description of a chemical scaffold.

Molecular dynamics were undertaken using AMBER v11[329] to rationalize the stabilisation of the telomeric DNA/RNA hybrid duplex sequence 5′-TTA-GGG-TTA-GGG-TTT-TTT- CCC-UAA-CCC-UAA-3′ by compounds **3.4** and **3.6** (**Figure 3B.3**) compared to other members of library-1 (for example, compound **3.3**). The minor groove of the hybrid sequence was chosen as the likely point of interaction for the molecules due to the similarity in structure of the molecules to the minor groove binders propamidine[330] and pentamidine[331]. A previous study

[2] The molecular modelling and molecular dynamics studies described here were carried out by Dr Paul J M Jackson and Mr Meir Touitou within the Institute of Pharmaceutical Sciences (IPS) at King's College London.

has suggested that DNA/RNA hybrid duplex-interactive molecules may act at least partly through an intercalation mechanism[192], so intercalation was also investigated as a potential mode of binding. In order to investigate their mechanisms of action, selected molecules in the series (for example, compounds **3.3**, **3.6** and **3.7** and **3.8**) were docked in the minor groove and simulations were carried out over a time period of 10 ns in implicit solvent.

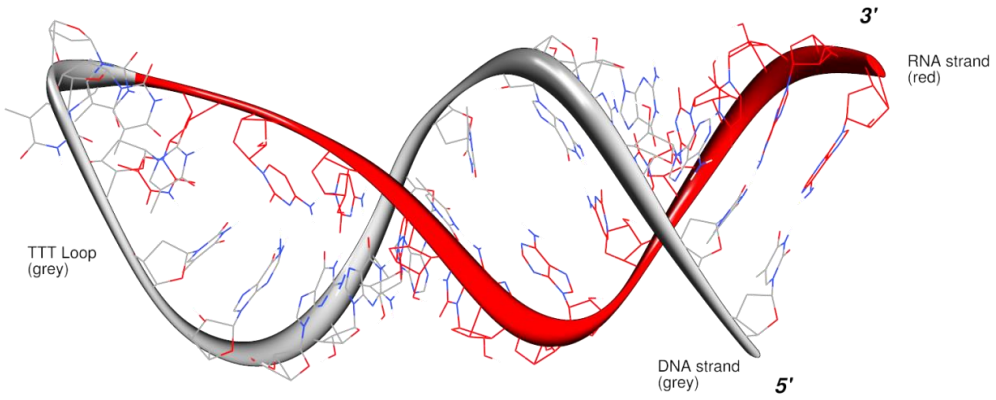


Figure 3B.3: Molecular model of the DNA/RNA hybrid duplex (5′-TTA-GGG-TTA-GGG-TTT-TTT- CCC-UAA-CCC-UAA-3′). The RNA strand is represented in red and the DNA strand in grey. DNA and RNA bases are indicated in stick form.

The results from simulations suggested that both compounds **3.4** and **3.6** (molecules with 9- and 11.5- fold selectivity for the telomeric DNA/RNA hybrid duplex over the cDRH sequence at concentrations of 1 μM, respectively) have a potential dual mechanism of action, whereby most of the molecule binds in the minor groove but both the benzimidazole moieties may concurrently intercalate into the DNA. This intercalation process is facilitated by the CH_2 group between the phenyl and benzimidazole moieties, which allows flexibility of the benzofused moiety and a favourable change in curvature of the ligand.

In the case of compound **3.6**, the flexible octamethylene linker appears to form favourable van der Waal's interactions with the central TA portion (*i.e.*, T8:A23, T9:A22), with one benzimidazole moiety intercalating between G10:C21 and G11:C20 and the second benzimidazole intercalating between T7:A24 and T8:A23. As such, the molecule spans four base-pairs (5'-**TTA-G**-3') (**Figure 3B.4**). Therefore, it is proposed that the synchronous minor

groove binding and intercalating ability of both **3.4** and **3.6** contribute significantly to their enhanced affinity for telomeric DNA/RNA hybrid duplex sequences.

To investigate further the interaction of this molecule set with cDRH, simulations in implicit solvent were also undertaken of compounds **3.3**, **3.6** and **3.7** interacting with a duplex DNA sequence (*i.e.*, 5′-TTA-GGG-TTA-GGG-TTT-TTT-CCC-TAA-CCC-TAA-3′) equivalent to the DNA/RNA hybrid duplex structure. In these simulations, **3.3**, **3.6** and **3.7** remained in the minor groove of the DNA for the duration of the simulation, occupying a similar orientation to pentamidine and propamidine and forming van der Waal's interactions with the DNA bases below. However, in these experiments, the benzimidazole moieties of the molecules did not intercalate into the DNA but instead oriented at angles perpendicular to the central methylene linker of the ligand (**Figure 3B.4**). This difference in interaction (*i.e.*, intercalation with the DRH sequence but not with the cDRH sequence) was further supported by free energy of binding calculations (kcal/mol) for each ligand in implicit solvent, which indicated a greater affinity of each of the library-1 molecules for the telomeric DNA/RNA hybrid duplex compared with the control cDRH sequence.

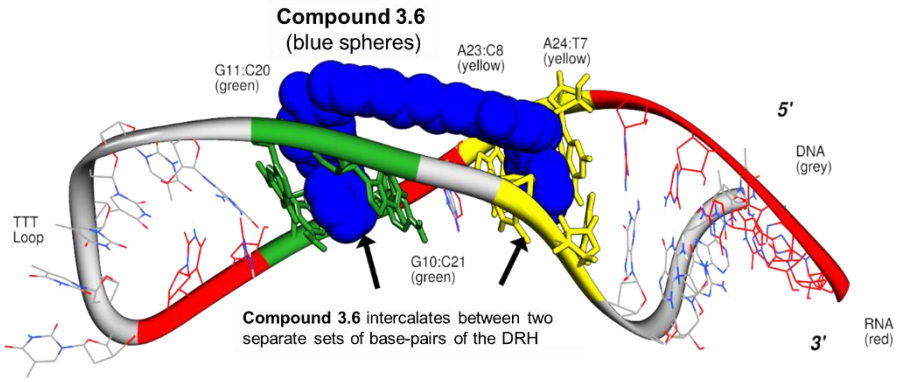


Figure 3B.4: Snapshot of a 10ns implicit solvent molecular dynamics simulation of compound **3.6** (blue spheres) interacting with DRH. Both benzimidazole moieties of the molecule intercalate between base pairs; one between G11:C20 and G10:C21 (green) and the second between A23:T8 and A24:T7 (both yellow).

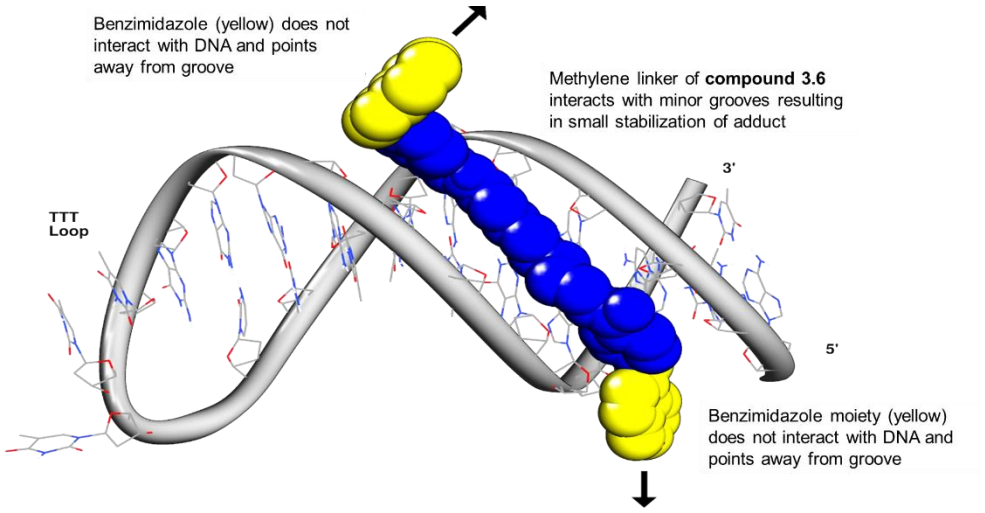


Figure 3B.5: Snapshot of a 10ns explicit solvent molecular dynamics simulation of compound **3.6** (blue and yellow spheres) interacting with the control duplex sequence, cDRH. Both benzimidazole portions (yellow) point away from the DNA groove and do not interact with the cDRH.

Simulations of **3.8** illustrate the molecule binding to DRH in a similar manner to that of compound **3.6**. The central methylene linker forms non-covalent interactions with the central tract of DNA (in this instance the sequence 5′-GTTA-3′), and forms a dual intercalation in the DRH structure (between G4:C27 and G5:C26 A9, and between nucleobases A9 and G10) (**Figure 3B.6**). It is likely that this intercalation contributes to the molecule's moderate observed stability of DRH (6.8 °C in FRET melting experiments). This feature, coupled with its ability to bind to a long (seven base-pair) sequence, may also explain the potent cytotoxicity observed for the molecule. Free energy of binding calculations also suggest that **3.4** possesses a preference for DRH compared with the control cDRH DNA sequence.

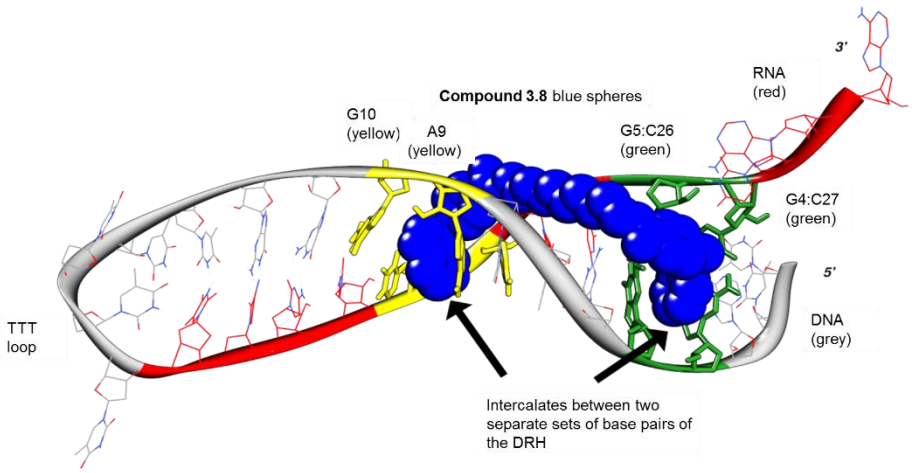


Figure 3B.6: Snapshot of a 10ns implicit solvent molecular dynamics simulation of **3.8** (blue spheres) with DRH. Both benzimidazole moieties of the molecule intercalate the sequence; one between G4:C27 and G5:C26 (green) and the second between the A9 and A10 nucleobases (yellow).

Further support for the superior interaction of library-1 molecules with the DRH sequence was evident from the fact that the molecules do not have the same curvature ("isohelicity") as DNA minor groove binders (for example netropsin[332] or pentamidine[331]), and the shallower topology of the telomeric DNA/RNA hybrid duplex may accommodate the less curved library-1 molecules to a greater extent. It follows that NSC-273829 is also less isohelical than the duplex DNA minor-groove binding agents (for example distamycin), and as such shows enhanced interaction with the telomeric DNA/RNA hybrid duplex compared with duplex DNA (**Figure 3B.7**).

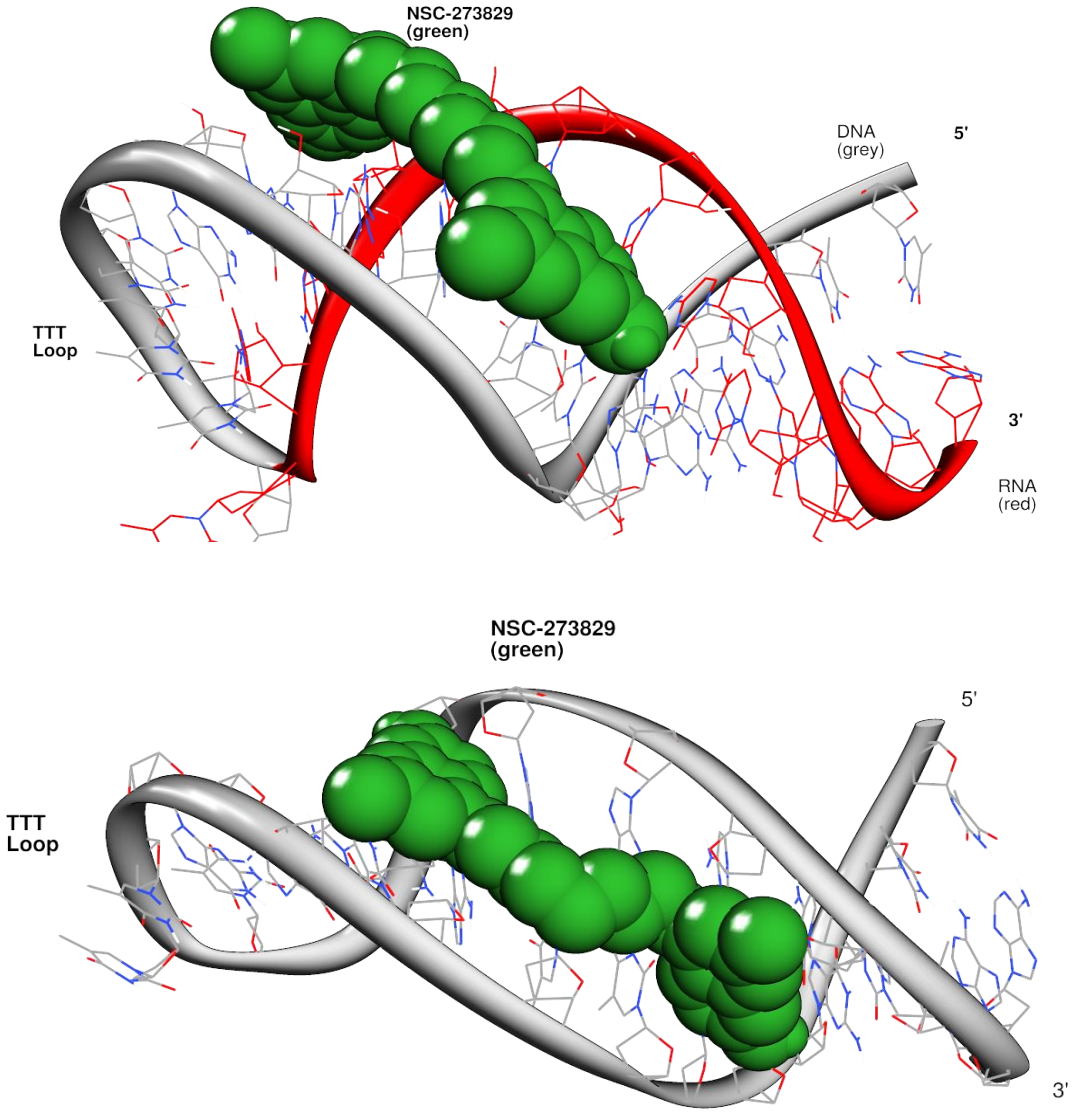


Figure 3B.7: Snapshot of a 10ns explicit solvent molecular dynamics simulation of NSC 273829 (green spheres) interacting with the DRH sequence (**top**, grey and red) and the cDRH (**bottom**, grey). The molecule is accommodated in each, but the shape-fit of the molecule to the DRH is more favourable due to the different topological features of the DRH structure.

The free energy of binding calculations correlated with the experimental results, and suggested that the library-1 molecules interact with the DRH sequence with a higher affinity compared with control hairpin duplex DNA (cDRH). It was also evident that compounds **3.4** and **3.6** (*i.e.*, those possessing the greatest stabilisation potential for DRH) contain an even number of methylenes (6- and 8-methylene spacers, respectively), whereas other molecules

in the series (*i.e.*, compounds **3.3**, **3.5** and **3.7**) contain an odd number of methylenes (5-, 7- and 9-methylene spacers, respectively). Therefore, it was hypothesised that there is a greater potential for dual intercalation in the case of ligands containing an even number of methylenes (*i.e.*, compound **3.4** and **3.6**), potentially explaining their higher affinity for telomeric DNA/RNA hybrid duplex sequences. This was supported by the molecular dynamics simulations, in which compound **3.6** was observed to be *bis*-intercalating with DNA (**Figure 3B.5**). This could be due to slight changes in the curvature of the molecules, or it is possible that an even-numbered linker may orient the benzimidazoles in a downward-facing direction, while an odd-numbered linker may facilitate the opposite.

Table 3B.5: Free energy of binding calculations (kcal/mol) of the intercalation of library-1 molecules and NSC 273829 with DRH and cDRH in explicit solvent.

Molecules	Implicit Solvent (PBSA)	
	DRH (kcal/mol)	cDRH (kcal/mol)
Compound 3.3	- 46.17	- 44.54
Compound 3.6	- 51.49	- 48.27
Compound 3.8	-53.62	-49.44
Compound 3.7	- 59.93	- 55.71
NSC-273829	- 29.76	- 35.59

Additionally, measurements of compound spacer length relating to the distance between the G10 and A23 nucleobases of the telomeric DNA/RNA hybrid duplex sequence were calculated through molecular modelling studies using VMD software. Using this approach, it was estimated that the methylene spacer between the phenyl-benzimidazole moieties of compound **3.6** was 15.64 Å while that between G1 and G5 of the DRH sequence was 13.13 Å. The distance of the methylene spacer between the quinoline moieties was also measured for NSC 273829 and found to be 9.97 Å, relatively shorter than the *bis*-intercalation distance (between G1 and G5) for the telomeric DNA/RNA hybrid duplex sequence. The flexibility of the methylene spacer between the two phenylene moieties added to the flexibility, allowing

compound **3.6** to bind in the minor groove of DRH, with most of the molecule bound in the minor groove and with the two benzimidazole moieties intercalated between nucleobases in the DRH sequence. However, changing the length of the methylene spacer and thus changing either the distance between the benzimidazole units or the type of spacer (for example, inducing rigidity in the case of **3.9**) caused either partial intercalation (for example, in the case of molecule **3.72**), weak interaction with the minor groove or no ligand-DNA interaction at all.

For example, in the case of compound **3.39** (library-3B), the methylene linker of the molecule interacts with the DRH sequence to some degree, but the fluorotetrahydroacridine substituents fail to intercalate into the DNA/RNA hybrid duplex structure, and instead form intra-molecular stacking interactions. The lack of intercalation of the terminal substituents is most likely due to the large nature of these substituents relative to the purine/pyrimidine bases (*i.e.*, see **Figure 3B.8**).

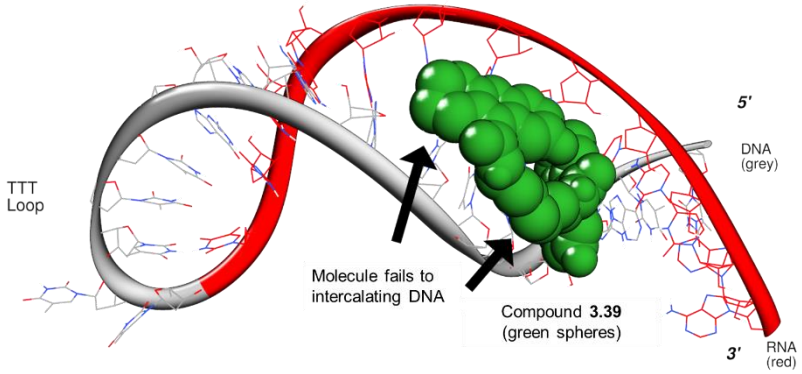


Figure 3B.8: Snapshot of a 10ns molecular dynamics simulation of compound **3.39** (green) in the minor groove of the DRH structure. The molecule does not intercalate into the DRH structure and fails to form any significant bonding.

Compound **3.22**, on the other hand, forms non-covalent interactions (for example, van der Waal's interactions) within the minor groove of DRH, but does not intercalate into the structure in any way (**Figure 3B.9**). Instead, the fluoromethylquiniline moieties of the ligand point along the floor of the minor groove of DRH and do not form H-bonds with the DNA/RNA bases. Presumably, this contributes to the lack of stabilisation observed in the FRET experiments,

and highlights the importance of intercalation as the primary mechanism for stabilisation of DRH.

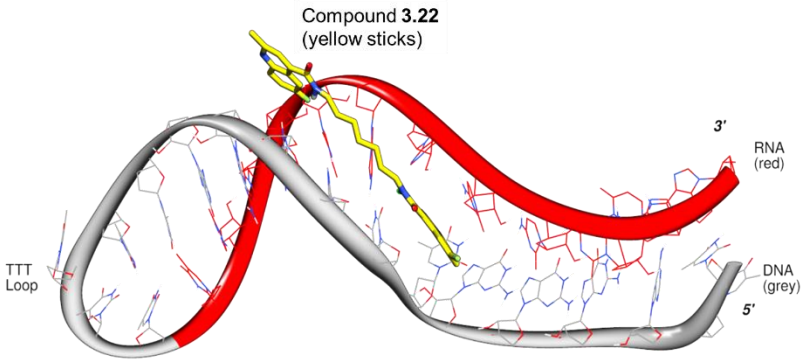


Figure 3B.9: Snapshot of a 10ns molecular dynamics simulation of compound **3.22** (yellow) seated in the minor groove of the DRH structure. The molecule does not intercalate within the DNA and fails to form any significant bonding interactions with the DRH structure.

Compound **3.10** was designed as an analogue of **3.6**, with a benzofused moiety linked *via* a biphenyl to the central linker. In this instance, the central linker was composed of a number of ring structures, and this ligand was designed to establish the importance of the flexible methylene linker in **3.6**. Simulations suggested that it is likely that compound **3.10** only part-intercalates into the DRH structure, and this is reflected in the poor FRET stabilisation temperatures observed for this molecule (*i.e.*, ΔT_m values were 3.6 and 3.2 °C at 2 μM and 1 μM ligand concentrations, respectively). The central biphenyl linker rigidifies the ligand structure, thereby preventing it from intercalating into DNA (**Figure 3B.10**). Furthermore, the molecule does not form any electrostatic interactions (*e.g.*, H-bonds) with the DNA/RNA floor, suggesting that its shape and orientation in the minor groove are not conducive to DRH interaction.

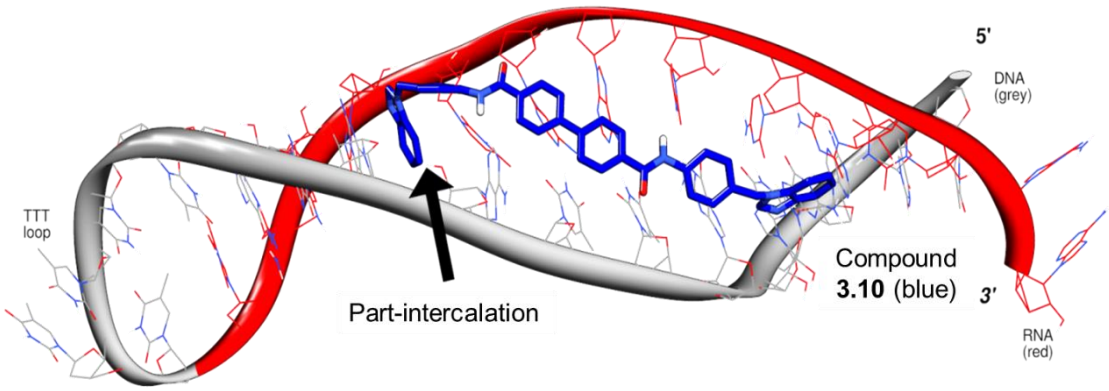


Figure 3B.10: Snapshot of a 10ns molecular dynamics simulation of compound **3.10** (blue) seated in the minor groove of the DRH structure. The molecule part-intercalates into the DRH structure but fails to form significant bonding interactions with the DRH structure.

Compound **3.9** has a similar type of structure to **3.10**, although the number of ring structures was decreased to investigate whether a reduction in the number of benzenes may induce more favourable interactivity with the DRH structure. Molecular dynamics simulations conducted over a 10ns time-period suggested that, in a similar manner to **3.10**, compound **3.9** produces few interactions with the minor groove floor, and does not intercalate with DRH in any way. The benzimidazoles at both terminals of the molecule point out of the DNA (*i.e.*, **Figure 3B.11**), and so are not capable of intercalation. Also, the reduction in size of the phenyl linker has little effect on the interaction of the molecule with DRH, and the molecular models show that the methylene linker between benzofused moieties is critical for DRH binding to occur.

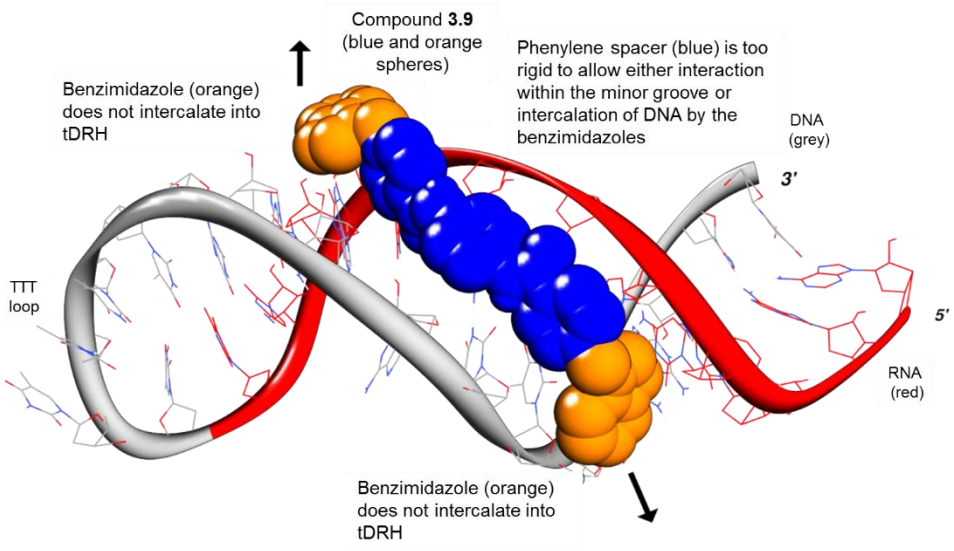


Figure 3B.11: Snapshot of a 10ns molecular dynamics simulation of compound **3.9** (blue) seated within the minor groove of the DRH structure. The molecule does not intercalate within the DRH structure, and fails to form any significant interactions with the DRH structure.

In an identical manner to **3.22**, both compounds **3.14** and **3.29** bind in the minor groove of DRH. However, they do not intercalate into the DRH structure, and do not form any significant interactions with the groove floor, and this was reflected in the poor DRH stabilisation temperatures observed in the FRET melting studies.

In summary, from the molecular modelling and molecular dynamics studies of the ligands synthesized to interact with the telomeric DNA/RNA hybrid duplex, it can be concluded that the molecular dynamics simulations of all libraries (library 1-4) provide insight into the binding mechanism of these ligands. It is evident from these simulations that a number of traits are critical for DRH binding. First, a methylene linker must be present in order for intercalation into the DRH structure to occur. Furthermore, linkers of even-numbered length stabilise DRH to a greater extent than odd-numbered linkers, and this occurs due to the orientation of the benzimidazole groups. An even-numbered linker will orient both benzimidazoles in a downward-facing direction, whereas an odd-numbered linker will result in only one benzimidazole moiety pointing downwards, and as a result will induce less stabilisation.

Secondly, a flexible linker between the benzimidazole units is critical for activity, and this is evident from simulations of compounds **3.9** and **3.10**, in which rigidifying the linker through the introduction of phenyls removes the molecule's ability to intercalate into DNA. Finally, the overall shape of the molecule is critical for effective stabilisation of DRH, and this is evident from the simulations of compounds **3.29** and **3.14** in which the molecules bind into the minor groove, but do not have the appropriate shape to intercalate.

In addition to modelling studies on the DRH sequence, further molecular modelling and molecular dynamics investigations were undertaken in an effort to rationalise the results of the FRET binding experiments undertaken on all the library molecules (except library-4, *bis*-benzofuran type compounds) with the F21T telomeric G-quadruplex. The F21T G-quadruplex (PDB ID: 3CDM[333]) was used as the DNA receptor, and the molecules of library-1 to library-3 were docked with the quadruplex using an established protocol optimised within our laboratories.

The FRET binding results suggested that molecules containing phenyl-benzimidazole moieties separated by a methylene (*i.e.*, compounds **3.3**, **3.4**, **3.5**, **3.6** and **3.7**) stabilise DNA to some degree, with stabilisation temperatures between 0.2 °C and 6.5 °C (for a 1 μM ligand concentration) observed for library-1. Docking studies showed that, although the molecules are large and do not have the required curvature to stabilise the quadruplex interface, some members of the series have the potential to *bis*-intercalate with the DNA quadruplex structure. For example, compound **3.5**, which has an F21T stabilisation temperature of 5.5 °C, was observed through the modelling to *bis*-intercalate with DNA with one benzimidazole interacting with the top face of the quadruplex and the second benzimidazole interacting with the bottom face (**Figure 3B.12**). Interestingly, a strong correlation (R^2 value of 0.77) was observed between the FRET melting temperatures and the GBSA (in general, GBSA refers to a Generalized Born model augmented with the hydrophobic solvent accessible surface area (SA), which is the most frequently used implicit solvent model combination) docking scores, suggesting that the docking results are an accurate representation of the ligand-DNA complex.

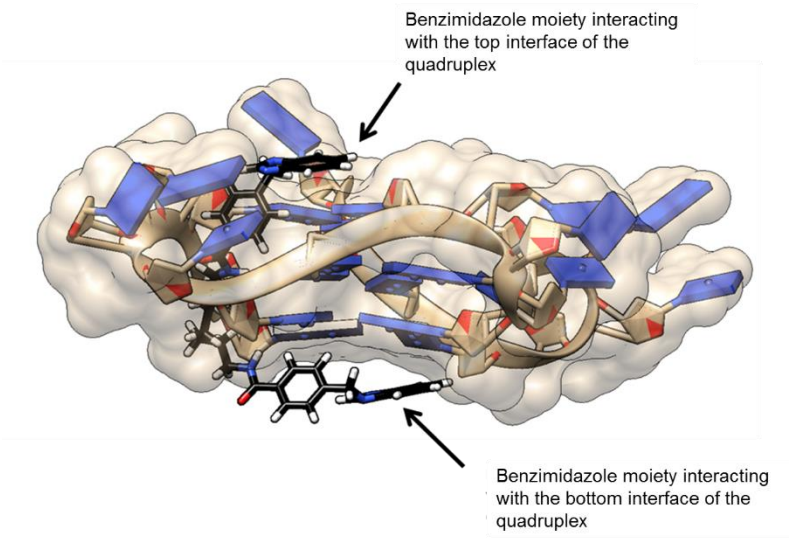


Figure 3B.12: Image of the highest scoring (GBSA) docked pose of compound **3.5** (black) interacting with the F21T quadruplex (PDB ID: 3CDM) (blue and white). The benzimidazole moieties interact with the top and bottom faces of the quadruplex through van der Waal's interactions, thus providing stabilisation.

Docking studies were also undertaken on one member of each of library, for example library-2A, -2B, -2C, 3A and -3B (see **Table 3B.6**), in an effort to rationalise their lack of affinity for the quadruplex structure. Study design was based on the library-1 molecule set, where compound **3.6** (with a carbon linker number of eight) was shown to interact with the highest affinity for telomeric quadruplex DNA. Molecules corresponding to compound **3.6** (*i.e.*, with a linker number or span equivalent to compound **3.6**) were chosen from each of the libraries (except library-4), and were docked on the F21T quadruplex.

Although the binding scores were reasonably favourable for some examples (*e.g.*, compound **3.9**) (see **Table 3B.6**), visual analysis of the models suggested that little stabilisation of the G-quadruplex structure was occurring. For example, in the case of compound **3.9**, **Figure 3B.13**, the molecule did not either stack on the quadruplex interface or form van der Waal's interactions with guanine bases. The shape, size, curvature and length of the molecule were not conducive to quadruplex interaction as it had little similarity to known quadruplex stabilisers (*e.g.*, the biaryl polyamides). Furthermore, the docked orientation of the ligand was not at the

centre of the quadruplex, and as such would not be expected to stabilise the telomeric quadruplex to a great extent.

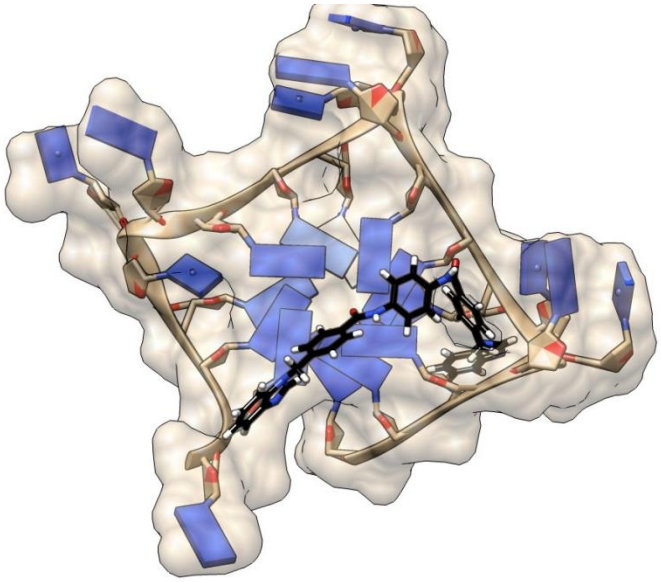


Figure 3B.13: Highest scoring (GBSA) docked pose of compound **3.9** (black) interacting with the F21T quadruplex (PDB ID: 3CDM) (blue and white). The molecule does not form stacking interactions with the quadruplex due to a poor shape-fit.

The FRET binding results suggest that library-3B molecules possess some degree of binding affinity for the quadruplex structure (for example, ΔT_m of compound **3.39** = 4.5 °C at 1 μM concentration), and this is most likely due to one of the fluorotetrahydroquinoline substituents (*i.e.*, library-3B) interacting with guanine bases at the quadruplex interface, thereby resulting in moderate stabilisation. Docking results support this hypothesis, and illustrate one half of the dimer forming van der Waal's interactions with the interface, but with the second half of the dimer failing to interact (**Figure 3B.14**).

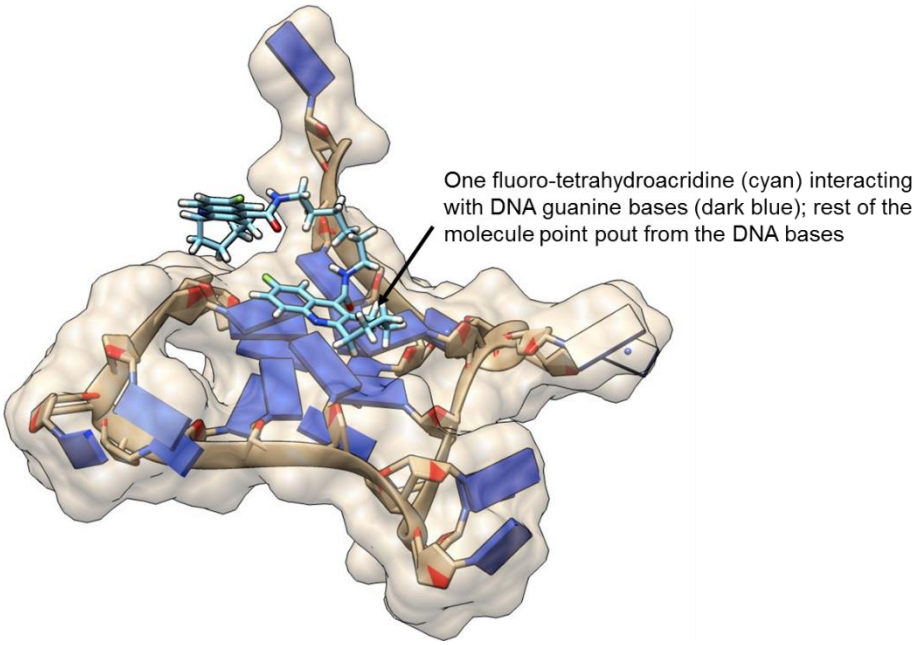
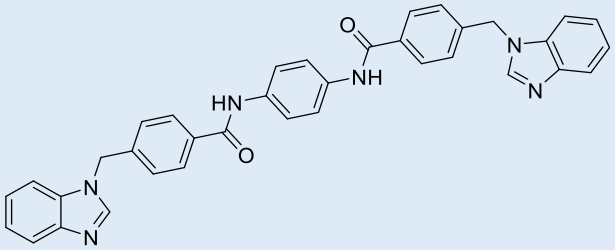
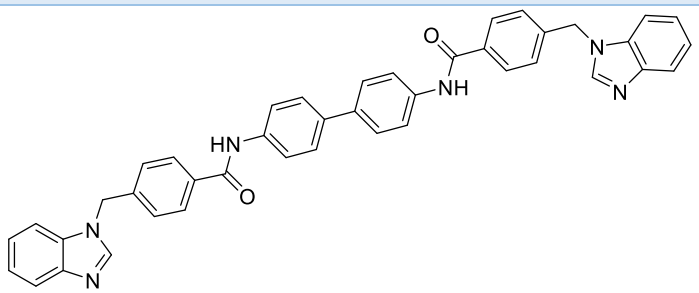
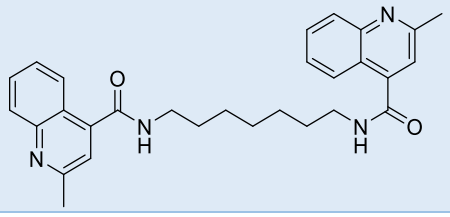
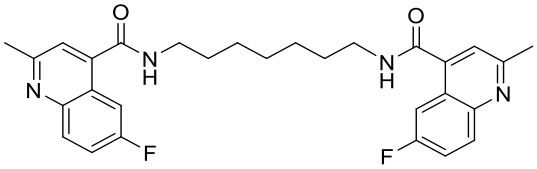
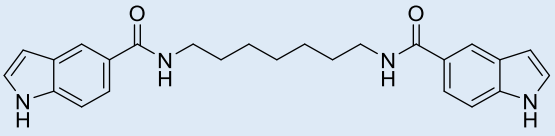
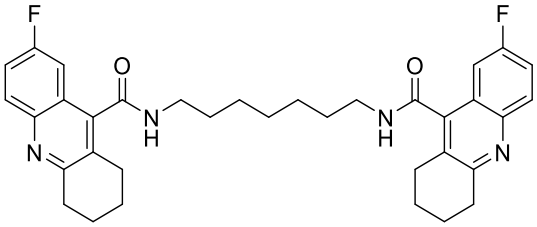
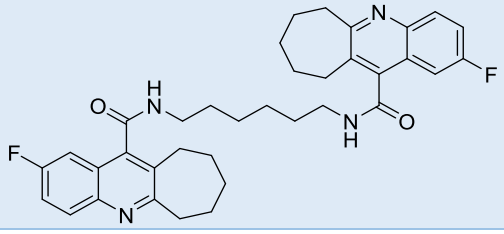


Figure 3B.14: Highest scoring (GBSA) docked pose of compound **3.39** (cyan) interacting with the F21T quadruplex (PDB ID: 3CDM) (blue and white). One fluoro-tetrahydroacridine moiety of the dimer forms stacking interactions with the DNA bases (blue nucleotides), whilst the second moiety fails to interact.

All other libraries were observed to stabilise the G-quadruplex by 2-3 °C in the FRET studies. Such stabilisation is likely to occur *via* non-specific interactions, and this was reflected in docking results. The results for compounds **3.14**, **3.22** and **3.29** docked on the F21T quadruplex illustrate this, and show the molecules binding non-specifically to the side of the quadruplex, in a similar manner to compound **3.9** (**Figure 3B.13**).

Table 3B.6: Docking scores (GBSA) of selected members from libraries 1-3 interacting with the F21T quadruplex (PDB ID: 3CDM).

Library	CODE	STRUCTURE	GBSA SCORE (kcal/mol)
Library-1	**3.9**		-68.86
	3.10		-45.03
Library-2A	**3.14**		-59.94
Library-2B	**3.22**		-43.53
Library-2C	**3.29**		-64.23
Library-3A	**3.34**		-49.40
Library-3B	**3.39**		-49.36

3B.3 Circular Dichroism (CD) Analysis

There is very little information available on CD data relating to DNA/RNA hybrid duplexes, especially those originating in the telomeric region of DNA. It is established that DNA/RNA hybrid duplexes can form an A-form DNA-type conformation and winding[334], and in 2012 Xu and colleagues reported that the hTelo DNA sequence can form a hybrid G-quadruplex within living cells in the presence of a complementary RNA sequence[297]. These researchers reported that the CD spectrum of telomeric DNA shows a positive CD signal at 290 nm and a negative signal at 265 nm (in the presence of Na^+). However, upon the addition of complementary RNA to the telomeric DNA, the peak at 290 nm disappears and is replaced by a strong positive peak at 260 nm and also shows a negative signal at around 240 nm. This change is similar to that observed for the transformation of antiparallel DNA to a parallel G-quadruplex. Moreover, a CD spectrum for RNA/DNA hybrid sequence (*i.e.*, poly(rA)-poly(dT)) also depicted (**Figure 2.10**) a positive CD signal at around 260 nm and a sharp negative CD signal at around 248 nm.[185] Thus it can be postulated that the hybrid sequence can show a negative signal at around 240 nm, which was identical to our CD spectroscopic analysis (**Figure 3B.16**).

Furthermore, Vorlíčková and co-workers in 2012 reported different CD spectra for various conformations of DNA structures, and for a DNA/RNA hybrid structure.[335] In their study, they used a DNA/RNA hybrid duplex in which the RNA sequence was r(GCGGCGACUGGUGAGUACGC), and the CD spectrum showed a strong positive CD signal at 275 nm and a negative one at 210 nm.

In this project, a hairpin-structured telomeric DNA/RNA hybrid duplex (DRH) (**Figure 3B.15 a**) was used to carry out the CD analysis, identical in structure to the DRH sequence used in the FRET assay, but without the labels and also the molecular modelling and molecular dynamics studies. A control DNA/DNA duplex (cDRH) was also used in the CD analysis, where the RNA nucleotides of DRH were replaced by DNA bases (**Figure 3B.15 b**).

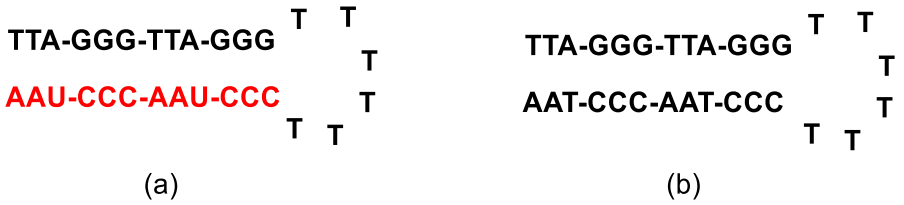


Figure 3B.15: Hairpin structures of (**a**) the telomeric DNA/RNA hybrid duplex used in the CD studies (**red** indicates the RNA nucleotides) and (**b**) the control cDRH sequence.

The CD spectrum of the hybrid duplex sequence (DRH) in Tris-HCl (50 mM) was initially measured and gave an intense positive CD signal at around 271 nm and a small negative CD signal at around 237 nm (**Figure 3B.15 a**). A CD spectrum of cDRH was also measured and showed a strong positive CD signal at around 269 nm and a negative one at around 241 nm (**Figure 3B.15 b**).

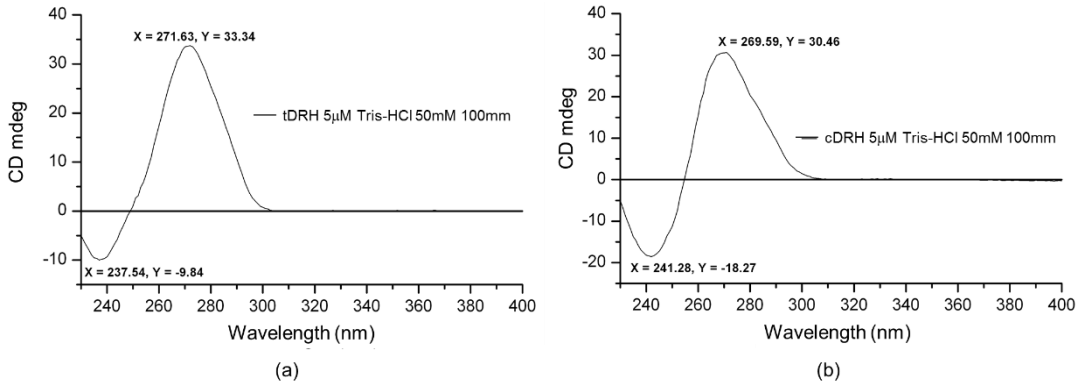


Figure 3B.16: CD spectrum of (a) the telomeric DNA/RNA hybrid duplex sequence (DRH), and (b) the control DNA duplex (cDRH).

Fourteen selected compounds were evaluated for their ability to interact with the telomeric DNA/RNA hybrid duplex and control cDRH sequences based on changes to the CD spectra. These fourteen molecules were carefully chosen on the basis of the FRET melting assay results and the molecular modelling and molecular dynamics studies. On this basis, seven *bis*-benzimidazole-type molecules were chosen from library-1 (*i.e.*, **3.4**, **3.5**, **3.6**, **3.7**, **3.8**, **3.9** and **3.10**), two *bis*-benzofuran-type molecules from library-4 (*i.e.*, **3.69** and **3.71**) and one molecule

from each of the other libraries (*i.e.*, **3.14**, **3.23, 3.28**, **3.35** and **3.40**). These molecules were chosen, based on previous FRET and modelling results, since they contained either seven-methylene or eight-methylene spacers between their two side arms.

The *bis*-benzimidazole compounds, **3.5**, **3.6**, **3.7** and **3.8**, showed concentration-dependent red shifts in the positive CD signal of between 1-4 nm. In the first instance, compound **3.5** in a 1:1 equivalent concentration with DRH (5 μM) gave a red shift of 1.2 nm for the positive signal at around 273 nm (**Figure 3B.17 a**). Additionally, the initial concentration of **3.5** (5 μM) gave a hypochromic shift in the positive signal (*i.e.*, ΔCDmdeg = 1.22). However, there was no noticeable change in the negative CD signal. On increasing the ligand concentrations (2-5 equivalents with DRH), the positive CD signal shifted further to the right and gave a red shift (at around 274 nm) at the highest ligand/DNA ratio (1:5). Moreover, the highest ligand concentration (*i.e.*, 25 μM) caused the maximum change in intensity of the CD signal (*i.e.*, ΔCDmdeg = 7.29) in the positive region. There was only a very small change in the intensity of the CD signal in the negative region without any blue and/or red shift. Compound **3.5** was also evaluated with the cDRH sequence, but no ligand-DNA interaction was observed up to 1-5 equivalents of drug (**Figure 3B.17 b**), confirming the selectivity of the *bis*-benzimidazole molecule for the DNA/RNA hybrid duplex.

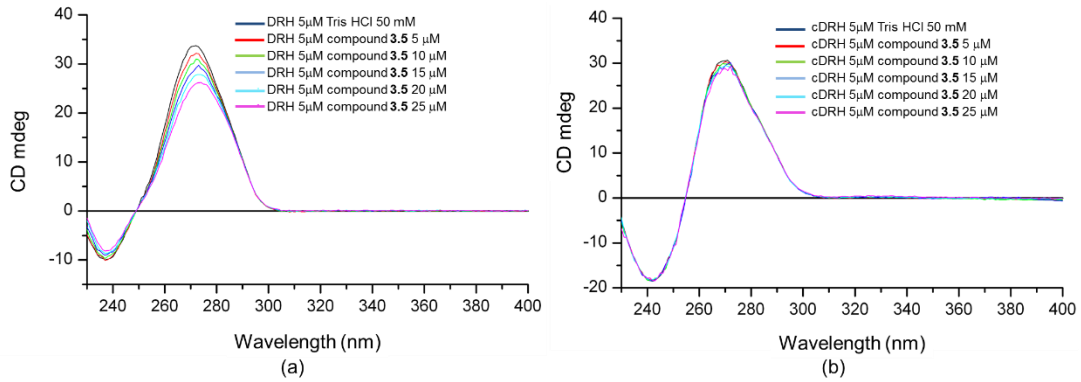


Figure 3B.17: CD spectrum of compound **3.5** with (a) DRH (5 μM) and (b) cDRH in Tris buffer (pH 7.4) at 0-5 equivalents ligand concentration (path length 10 mm).

In the case of compound **3.6**, the shifting of the CD signal in the positive region was more pronounced (**Figure 3B.18 a**). CD signals were observed at around 272.2, 273.0, 274.5, 275.0 and 275.5 nm (red shifts) for compound **3.6** with 1, 2, 3, 4 and 5 equivalents of ligand-DNA,

respectively. In addition to the red shift, there were hypochromic shifts similar to those observed for **3.5**. The highest concentration of **3.6** (*i.e.*, 25 μM) lowered the positive peak intensity by up to 24.11 CD mdeg (where, the ΔCDmdeg = 9.23). Compound **3.6** also shifted the negative CD signal, where the maximum red shift (238.02 nm) was observed at the 5 equivalent ligand concentration (*i.e.*, at 25 μM). All of these shifts and changes in intensity suggested that compound **3.6** was interacting with the DNA/RNA hybrid duplex sequence. This also supported data from the FRET melting assay which showed that compound **3.6** stabilized the DRH sequence by 7.2 °C at a 1 μM concentration. Similarly, the unchanged CD spectrum on addition of **3.6** to the cDRH sequence (**Figure 3B.18 b**) reflected the FRET melting results (*i.e.*, DRH/cDRH selectivity = 9.0 at a 1 μM ligand concentration).

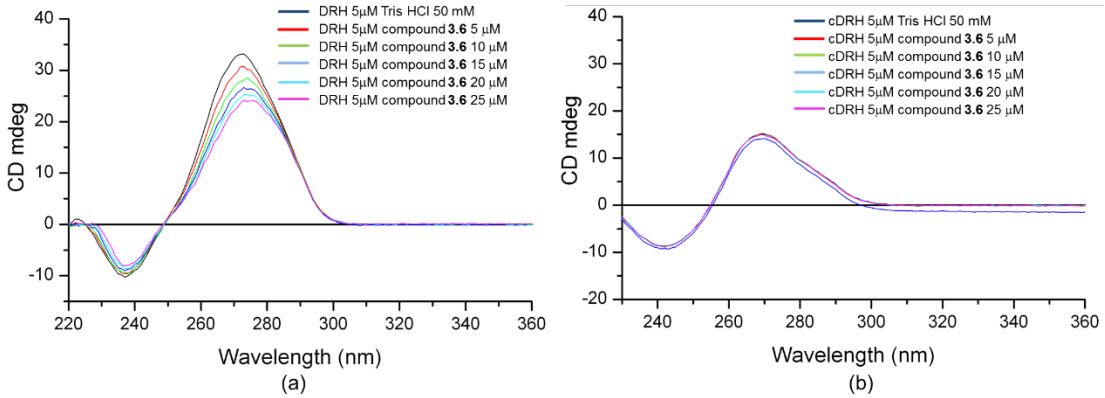


Figure 3B.18: CD spectrum of compound **3.6** with (a) DRH (5 μM) and (b) cDRH in Tris buffer (pH 7.4) at 0-5 equivalents ligand concentration (path length 10 mm).

Compound **3.8** also imposes both the red shift and hypochromic shifts on the positive CD signal. 1, 2, 3, 4 and 5 equivalents of **3.8** caused shifts in the positive signal by 1.46, 2.72, 2.99, 3.91 and 4.37 nm, respectively, towards the right of the spectrum, with the furthest signal observed at around 276 nm for the highest ligand-DNA ratio (**Figure 3B.19 a**). ΔCDmdeg changes of 3.56 and 11.56 were observed for the lowest (*i.e.*, 1 eq) and highest (*i.e.*, 5 eq) ligand concentrations, respectively. However, for the negative CD signal, there was little change with only 0.54 nm shifts observed for the highest ligand concentration. In this case, the CD data also correlated with the FRET melting results for **3.8** (*i.e.*, ΔT_m = 6.8 °C,

DRH/cDRH selectivity = 7.6 at a 1 μM ligand concentration). According to **Figure 3B.19 b**, no interaction was observed between the ligand and cDRH, confirming the molecule's selectivity towards the DRH sequence.

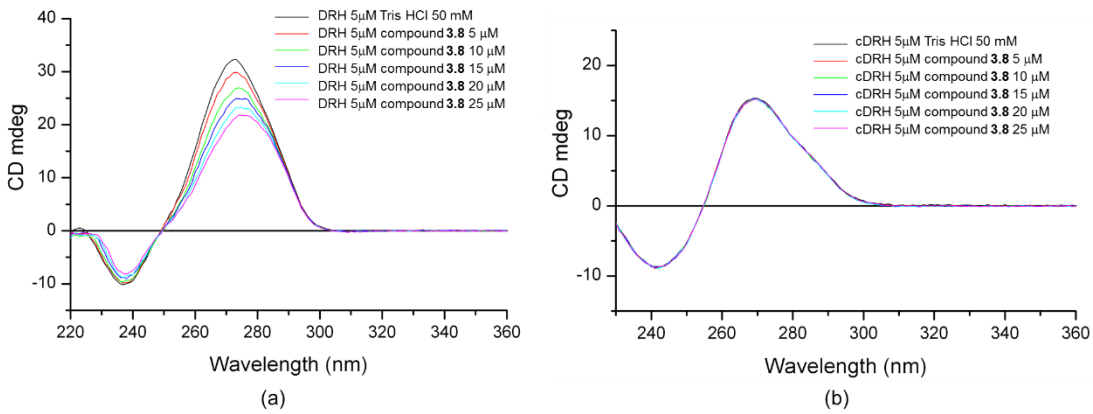


Figure 3B.19: CD spectrum of compound **3.8** with (a) DRH (5 μM) and (b) cDRH in Tris buffer (pH 7.4) at 0-5 equivalents ligand concentration (path length 10 mm).

Compound **3.7** showed selective interaction with the DRH sequence compared to the cDRH sequence according to the CD analysis (**Figure 3B.20**). The compound shifted the CD signal to around 277 nm and 272 nm for 5 and 1 equivalents of ligand concentration, respectively. Furthermore, **3.7** also caused a hypochromic shift of up to 12.83 mdeg for a 25 μM ligand concentration. However, no changes were seen with the cDRH sequence, confirming the lack of interactions observed in the FRET experiments.

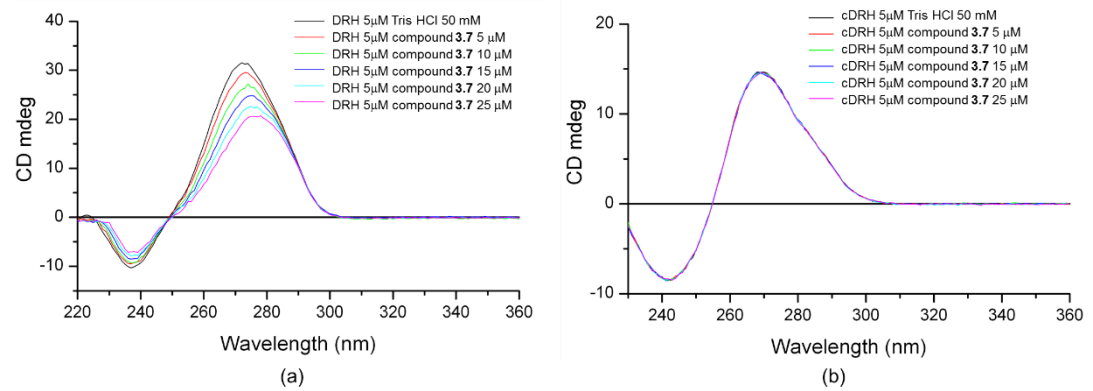


Figure 3B.20: CD spectrum of compound **3.7** with (a) DRH (5 μM) and (b) cDRH in Tris buffer (pH 7.4) at 0-5 equivalents ligand concentration (path length 10 mm).

On the other hand, due to changes in methylene spacer length (**3.4**) and type (*e.g.*, phenylene and biphenylene) (**3.9** and **3.10**), those compounds did not show any affinity for the DRH sequence.

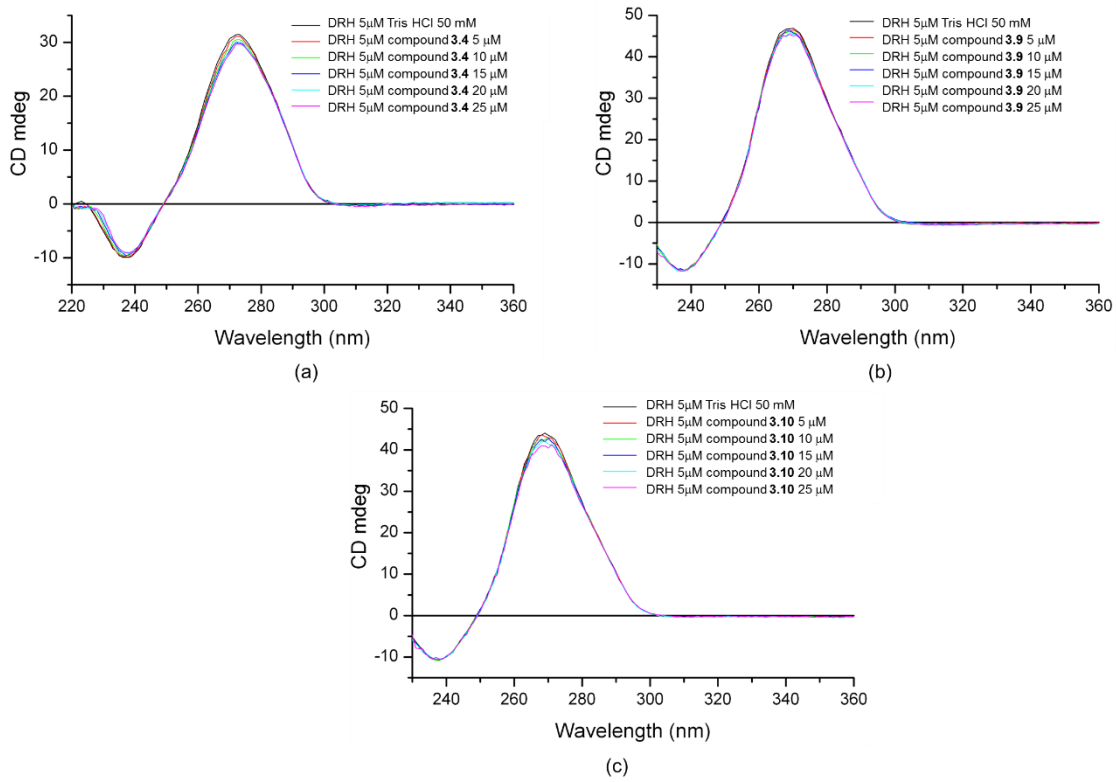


Figure 3B.21: CD spectrum of compounds (a) **3.4**, (b) **3.9** and (c) **3.10** interacting with the DNA/RNA hybrid duplex sequence (5 μM) in Tris buffer (pH 7.4) at 0-5 equivalents ligand concentration (path length 10 mm).

Molecular modelling and molecular dynamics studies suggested that such *bis*-benzimidazole-type molecules may interact with the telomeric DNA/RNA hybrid duplex via a dual mechanism involving both intercalation of the benzimidazole moieties between nucleobases and minor groove binding by the methylene spacer unit. However, MD studies on compound **3.6** confirmed that both benzimidazole units of the molecule can potentially intercalate, with one benzimidazole moiety inserting between G11:C20 and G10:C21 and the second between A23:T8 and A24:T7 (**Figure 3B.4**). Further molecular modelling and molecular dynamics studies established that certain lengths of the spacer unit can allow the molecule to intercalate via its benzimidazole moieties. For **3.6**, it was estimated that the methylene spacer between the phenyl-benzimidazole moieties was 15.64 Å, while the distance between G1 and G5 in

DRH was calculated to be 13.13 Å. Moreover, compounds **3.9** and **3.10** were designed to have the flexible methylene spacer replaced with planar phenylene and biphenylene spacers, respectively. Thus, it can be hypothesized that the rigidity of the planar spacer system prevents the molecules from having the appropriate curvature to fit into the DRH sequence.

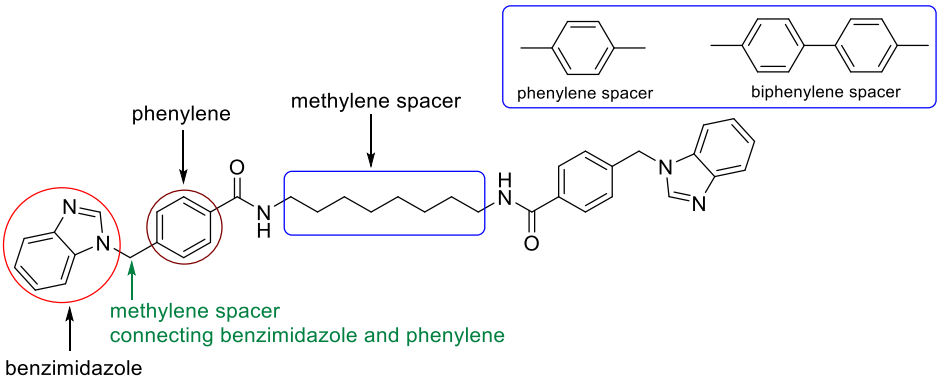


Figure 3B.22: Highlighted features of the *bis*-benzimidazole molecules showing compound **3.6** as an example. The phenylene and biphenylene spacers are present in **3.9** and **3.10**, respectively.

One *bis*-(2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxamide type molecule (**3.40**) with a seven-methylene spacer was selected from library-3B for CD analysis. Though the molecule did not show very good selectivity for the DRH sequence compared with the cDRH in the FRET melting assay, the compound gave noticeable CD signal changes in both the positive and negative regions of the spectrum (**Figure 3B.23**). The compound gave a ΔT_m value of 2.3 °C at a 1 μM ligand concentration with a DRH/cDRH selectivity of 1.9, which did not correspond well with the CD results. The CD signal showed a 3.05 nm red shift for the positive signal with 5 equivalents of ligand (*i.e.*, 25 μM) for the DRH sequence, but in the negative region of the spectrum there was little change. The CD signal shifted toward the positive region according to the ligand concentration. Although the CD signal gave a positive indication of ligand/DNA interaction, molecular modelling and molecular dynamics studies on compound **3.33** (a similar library molecule) established that the methylene linker of the molecule may bind to the DRH sequence to some degree, but the fluorotetrahydroacridine substituents fail to intercalate into the DNA/RNA hybrid duplex structure, instead forming intramolecular stacking interactions. Thus, it can be concluded that the changes in CD signals

for compound **3.40** might be due only to the methylene spacer, which explains why the compound is not very selective for the DNA/RNA hybrid duplex over the cDRH sequence.

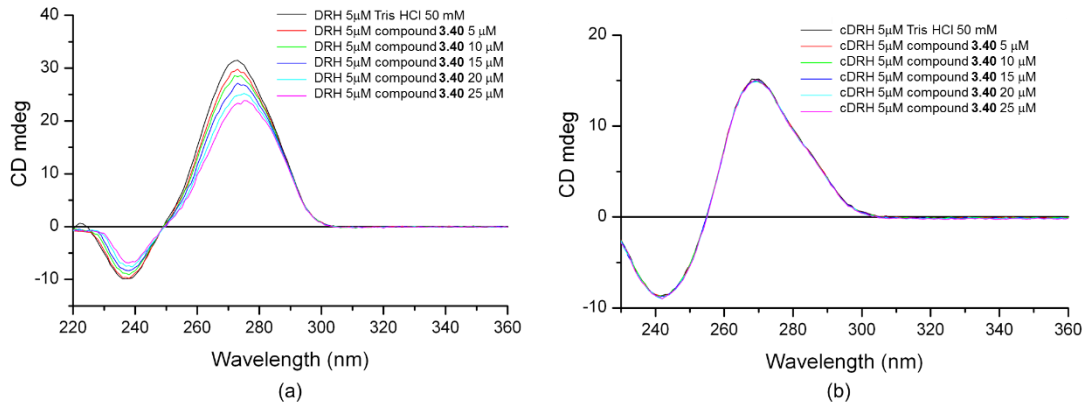


Figure 3B.23: CD spectrum of compound **3.40** with (a) DRH (5 μM) and (b) cDRH in Tris buffer (pH 7.4) at 0-5 equivalents ligand concentration (path length 10 mm).

In this study, the *bis*-indole ligand **3.28**, the *bis*-acridine ligand **3.35**, the *bis*-(6-fluoro-2-methylquinoline) ligand **3.23**, the *bis*-quinoline ligand **3.14** and the *bis*-benzofuran ligands **3.69** and **3.71** were also studied for interaction with the telomeric DNA/RNA hybrid duplex sequence (**Figure 3B.24**) and cDRH sequence by CD. However, none of these ligands induced any changes (red/blue shift) in the CD signals for the DRH sequence with 0-5 equivalents ligand-DNA concentration. However, a very small hypochromic shift in the positive CD signal occurred with the increase in ligand concentration, which was between 0-2 CD mdeg. Not surprisingly, all the ligands were inactive in the FRET melting assay, and no ligand-DNA interactions were obvious from the molecular modelling and molecular dynamics studies. According to the MD simulations, the terminal planar structures of each of the different molecule types (*i.e.*, *bis*-indole, *bis*-acridine, *bis*-(6-fluoro-2-methylquinoline) and *bis*-quinoline) were connected through seven- and/or eight-methylene spacers through amide bonding, thus allowing the molecules to interact with the minor groove of the DRH sequence to some extent but not facilitating intercalation (*i.e.*, the planar terminal side arms were either pointing out or only partially intercalating between nucleobases).

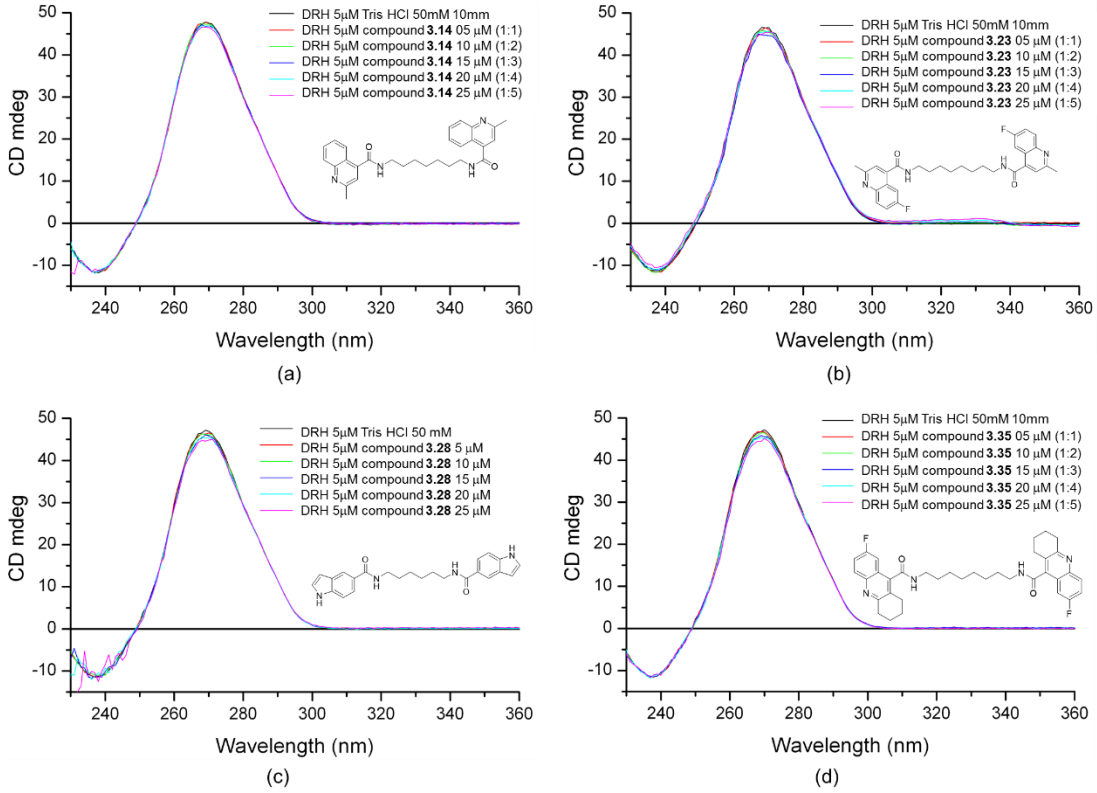


Figure 3B.24: CD spectrum of compounds (a) **3.14**, (b) **3.23**, (c) **3.28** and (d) **3.35** intercalating with DRH (5 μM) in Tris buffer (pH 7.4) at 0-5 equivalents ligand concentration (path length 10 mm).

As hypothesized from the molecular modelling and molecular dynamics studies, a ligand needs to be curved and crescent-shaped to interact with the DNA/RNA hybrid duplex, an observation established from the MD simulations of library-1 molecules, especially compound **3.6**. To confirm this hypothesis, two more molecules were synthesized within library-1 by incorporating phenylene and biphenylene spacers instead of the methylene spacer as in compounds **3.9** and **3.10**, respectively. In the CD analysis, no changes in the CD signal were observed, thus confirming the lack of affinity of the molecules for the DNA/RNA hybrid duplex (**Figure 3B.25**). This CD result correlated with the FRET melting results in which compounds **3.9** and **3.10** gave ΔT_m values of 0.5 and 3.2 °C at 1 μM ligand concentrations, respectively, with DRH/cDRH selectivity of 1.7-fold and 5.3-fold, respectively. Molecular modelling and molecular dynamics studies confirmed that molecules **3.9** and **3.10** are rigid and planar

structures compared with molecules containing methylene spacers. Thus they do not have the curved shape necessary to fit into the minor groove of the DRH sequence.

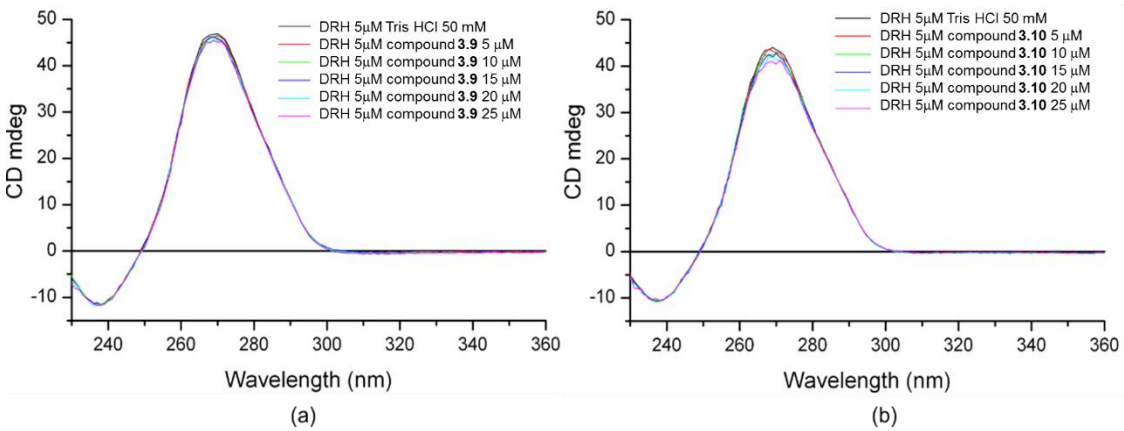


Figure 3B.25: CD spectrum of the interaction of compounds (a) **3.9** and (b) **3.10** with DRH (5 μM) in Tris buffer (pH 7.4) at 0-5 equivalents ligand concentration (path length 10 mm).

This study confirmed that the molecules need to have a curved shape, and have a degree of flexibility, as well as an appropriate length of the central linker in order to bind in the minor groove and intercalate between nucleobases (*i.e.*, compounds **3.2**-**3.7**). For this reason, the methylene spacer between the benzimidazole and phenylene moieties (*i.e.*, **Figure 3B.22**) is crucial. As some of the other molecules have methylene spacers of similar length to those in the library-1 molecules and also contain intercalating side arms, none of the molecules interacted with the telomeric DNA/RNA hybrid duplex.

Fourteen compounds were also evaluated for their interactions with the human telomeric G-quadruplex sequence (F21T; 5′-GGG-TTA-GGG-TTA-GGG-TTA-GGG-3′).

Human telomeric DNA includes several kilo-nucleobases of tandem duplications of the sequence 5'-TTAGGG that include a terminal single-stranded overhang of approximately 200 nucleotides. Intramolecular quadruplexes can adopt different topologies and can form lateral, edgewise and/or diagonal loops. CD experiments can identify the folding patterns and parallel or antiparallel configurations of G-quadruplexes.

It has been reported that parallel quadruplexes provide CD spectra with a 260 nm strong positive band and a 240 nm negative band, while an antiparallel topology provides a positive band at 290 nm.[7, 300, 336] The CD spectrum of the hTelo quadruplex sequence can show bands representing the presence of both parallel and antiparallel structures.[7] According to Paramasivan and co-workers, parallel quadruplexes display a positive CD signal at around 265 nm and a negative one at around 240 nm. On the other hand, the antiparallel quadruplexes provide a positive CD signal at around 295 nm and a negative signal at around 260 nm.[337] In 2012, Nicoludis and co-workers reported the CD spectra for the human telomeric sequence (d[$AG_3(TTAGGG)_3$]) in 5 mM K^+, 50 mM Na^+ and 100 mM Li^+ buffer at 25 °C.[338] The study showed two positive signals at around 290 and 240 nm and a sharp negative CD signal at around 260 nm in the presence of Na-buffer (**Figure 3B.26** a), which is identical to our study CD spectrum (**Figure 3B.26 b**).

In our study, the hTelo G-quadruplex showed the presence of mixed parallel and antiparallel topologies, with a 295 nm positive CD signal, a negative signal at around 265 nm and a further positive signal at around 245 nm.

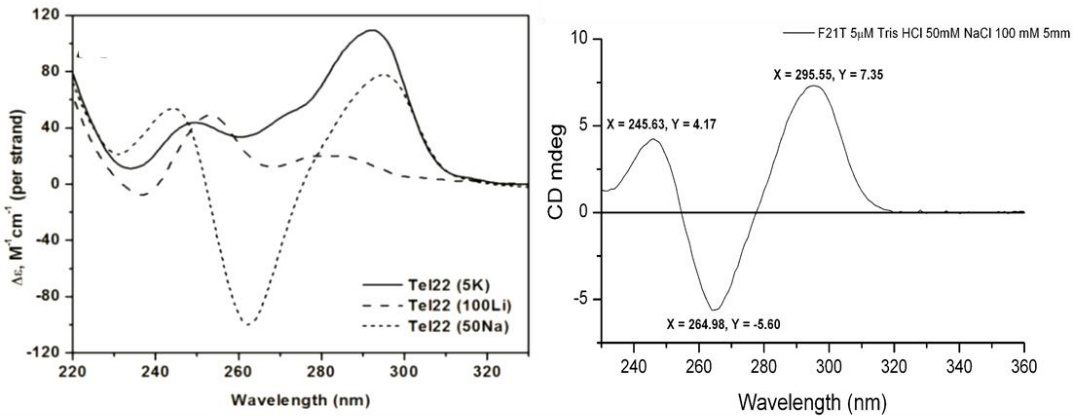


Figure 3B.26: CD spectrum of **(a)** human telomeric sequence ((d[$AG_3(TTAGGG)_3$])) in 50 mM Na-buffer and **(b)** telomeric G-quadruplex sequence (F21T) in Tris buffer (pH 7.4) in the presence of 100 mM NaCl.

A CD titration assay (**Figure 3B.27**) was carried out with 0-5 equivalents of fourteen selected ligands. However, none of the *bis*-benzimidazole-type molecules in library-1 caused any changes to the CD spectrum, even at 5 equivalents. These results contrasted with the FRET melting assay results (**Table 3B.2**) and molecular modelling and molecular dynamics studies (Section **3B.2.1** and **Figure 3B.12**) which suggested that the molecules do not interact with the telomeric G-quadruplex sequence. In addition to library-1 molecules, compounds **3.14** (*bis*-quinoline), **3.28** (*bis*-indole), **3.35** (*bis*-acridine type), **3.40** (*bis*- fluorotetrahydroquinoline) and **3.71** (*bis*-benzofuran) did not show any ligand-DNA interaction at 1-5 equivalents of ligand concentration.

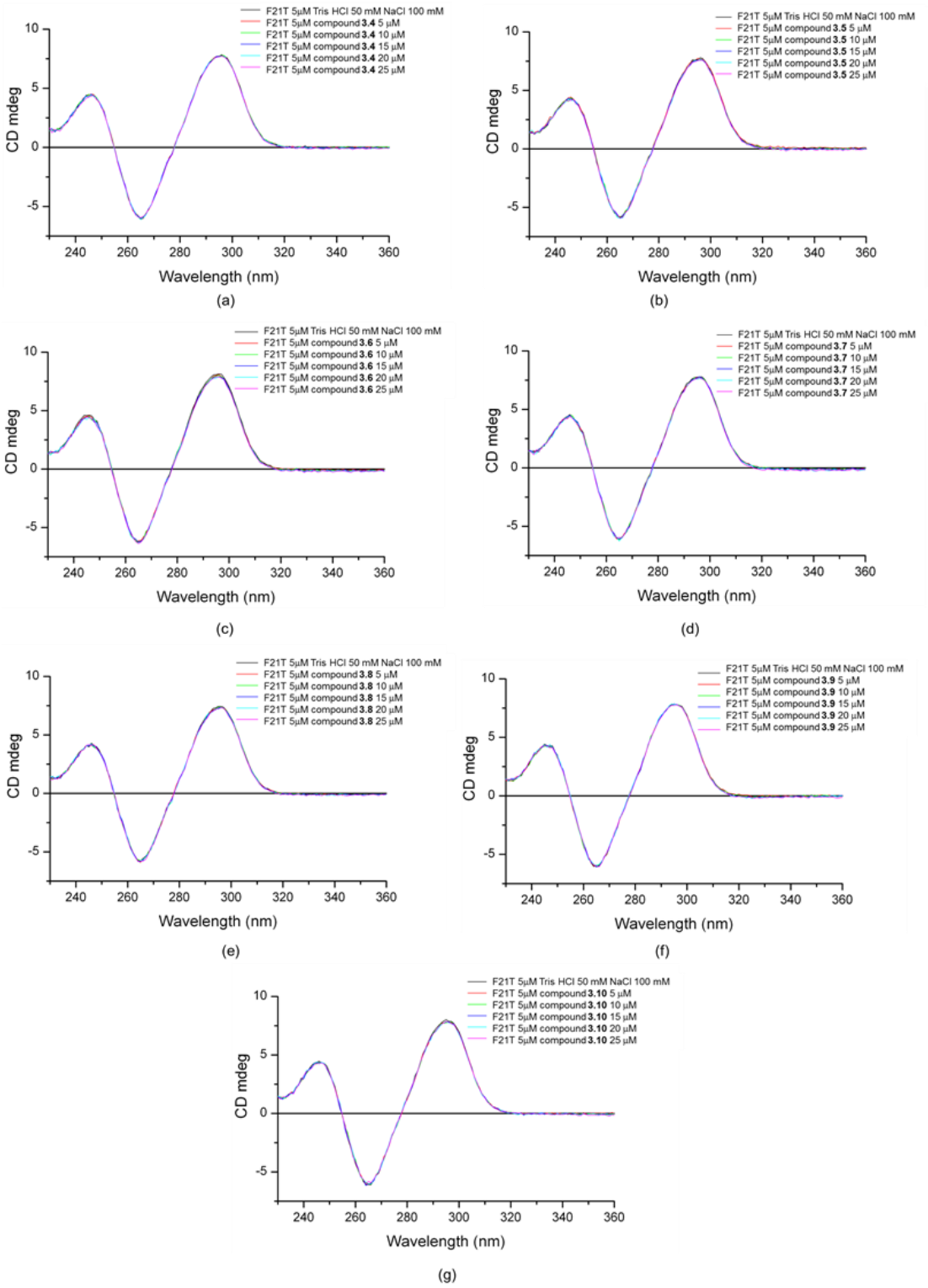


Figure 3B.27: CD spectrum of compounds (a) **3.4**, (b) **3.5**, (c) **3.6**, (d) **3.7**, (e) **3.8**, (f) **3.9** and (g) **3.10** interacting with the F21T sequence in Tris buffer (pH 7.39) in the presence of 100 mM NaCl at 0-5 equivalents ligand concentration (path length 10 mm).

Interestingly, compound **3.23** (*bis*-(6-fluoro-2-methylquinoline) and **3.69** (*bis*-benzofuran type, with a butyl-pyrrolidine tertiary amine tail) gave hypochromic shifts in the positive region of the CD spectrum. However, none of the molecules caused shifts of the CD signal toward the left or right of the x-axis (**Figure 3B.28**).

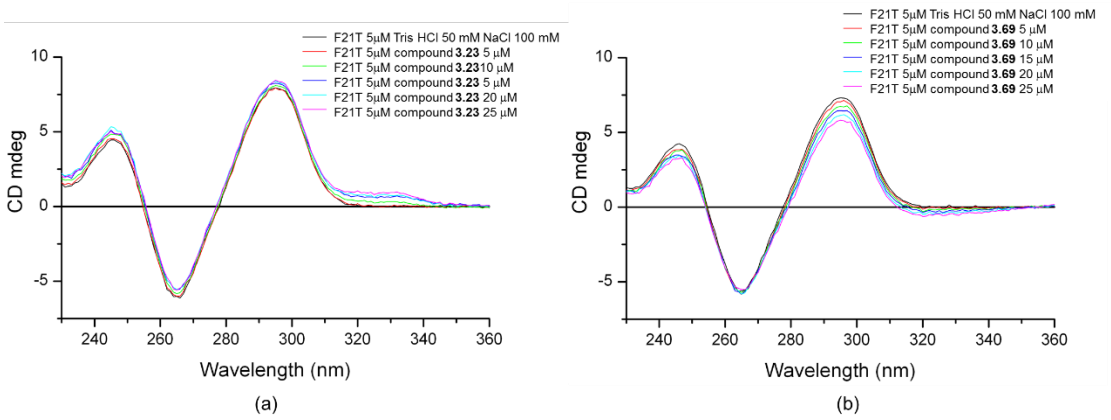


Figure 3B.28: CD spectrum of the interaction of compounds (a) **3.23** and (b) **3.69** with the telomeric G-quadruplex sequence (F21T) in Tris buffer (pH 7.39) in the presence of 100 mM NaCl at 0-5 equivalents ligand concentration (path length 10 mm).

In conclusion, the CD results presented here appear to confirm that the molecules in library-1 (except **3.2**, **3.9** and **3.10**) interact with the telomeric DNA/RNA hybrid duplex. Although CD data cannot confirm a molecule's mode of action as either stabilizing or destabilizing, based on the FRET melting results and molecular modelling studies, it can be hypothesized that the molecules stabilize the telomeric DNA/RNA hybrid duplex sequence.

3B.4 Cytotoxicity Test Using the MTT Assay

Novel potential anticancer agents are often initially evaluated for their cytotoxicity towards a panel of tumour cell lines. Although such *in vitro* cytotoxicity assays cannot predict either *in vivo* efficacy or toxicity, they can be valuable in the selection of novel compounds to progress to further pre-clinical studies. There are a number of different assays available, using both direct and indirect markers. In this project, the MTT assay was employed which uses the MTT

dye to make the cells visible, if still alive.

The characteristic parameter determined in this assay is the IC_{50} value which is the concentration at which 50% of cells die in the presence of the test molecule. One potential limit of the MTT assay is that it cannot distinguish between the cytocidal and cytostatic effects of a test molecule. However, the MTT assay is widely used due to its rapid, versatile and quantitative characteristics, and its reproducibility. The assay is based on the conversion of MTT to a purple formazan metabolite by mitochondrial enzymes within living cells[304], a process that is inhibited in the presence of a cytotoxic test molecule. This purple colour can be measured and quantitated.

The novel compounds synthesized in this project were evaluated in three human cancer cell lines: MDA-MB-231 (triple negative breast cancer cell line), HeLa (cervical cancer cell line) and NCI H1975 (non-small cell lung cancer cell line).

In total, forty-nine compounds were evaluated at a concentration of 25 μM by the MTT assay for their cytotoxic efficiency in the MDA-MB-231 cell line. Following this, 16 molecules with the lowest IC_{50} values were selected for further studies in the other two cancer cell lines (HeLa and NCI H1975). Compounds with IC_{50} values of 100 μM or above were considered as inactive. In this study, percentages are based on the assumption that the optical density of the cells without the compound present is 100%.

3B.4.1 MTT Assay Results with MDA-MB-231 Cell Line (Breast Cancer Cell Line)

In this study, the MDA-MB-231 cell line was seeded at 10,000 cells/well for the MTT assay. According to the IC_{50} concentrations described in **Table 3B.7**, the selected compounds from libraries -1 and -3B had greater cytotoxicity in the MDA-MB-231 cell line compared with the selected compounds from library-4.

Table 3B.7: Cytotoxic activities of compounds from library-1, library-3B and library-4 in the MDA-MB-231 cell line.

Library	Compound	IC_{50} Values (μM) * Mean (n=3)	R^2
Library-1	**3.4**	3.4 μM	0.4007
	3.5	4.2 μM	0.8256
	3.6	1.1 μM	0.6440
	3.7	9.2 μM	0.6933
	3.8	4.2 μM	0.8018
Library-3B	**3.39**	10.8 μM	0.8266
	3.42	3.6 μM	0.7530
	3.43	0.5 μM	0.7856
Library-4	**3.66**	> 100 μM	0.9029
	3.68	> 100 μM	0.7067
	3.69	> 100 μM	0.4862
	3.71	15.1 μM	0.9163
	3.73	> 100 μM	0.2428
	3.75	5.8 μM	0.6942
	3.77	5.6 μM	0.8815
	3.78	6.3 μM	0.8615

The IC_{50} values of most of the selected compounds from library-1 and library-3B were below 10 μM. Compound **3.6** had the lowest IC_{50} value of 1.1 μM from the library-1 molecules. It is noteworthy that this compound also showed the best selectivity for the telomeric DNA/RNA hybrid duplex sequence and telomeric G-quadruplex sequence compared with the respective control DNA sequences in the FRET melting assay at a 1 μM ligand concentration (**Table 3B.1**). However, compound **3.43** from library-3B had the lowest IC_{50} value (*i.e.*, 0.5 μM) among all tested compounds in the MDA-MB-231 cell line. In addition, the compound was the best fit

within the same oligonucleotide sequence in the molecular modeling and molecular dynamics studies (**Figure 3B.4**). It should also be noted that telomerase expression is higher in the MDA-MB-231 cell line due to upregulation of hTERT. Thus, the results for this breast cancer cell line with compound **3.6** suggest that the compound may be downregulating hTERT expression, thus leading to telomerase inhibition. Compounds **3.4**, **3.5**, **3.8** and **3.42** which have six-, seven-, ten-, and nine-methylene spacers, respectively, were also potent in this assay. However, compounds **3.4** and **3.8** had comparatively better activity than **3.5**, the rank order of which correlates with the hypothesis generated in the molecular modelling and molecular dynamics studies that molecules with even-numbered methylene spacers are more active than those with odd-numbers. Moreover, due to introducing rigidity in compounds **3.9** and **3.10** (though the C-number in the spacer unit is similar), these molecules lost activity in the MTT assay, also supporting the FRET results and the MD simulation studies that suggest the need for flexibility within the crescent-shaped molecules. Compounds **3.75**, **3.77** and **3.78** from library-4 molecules gave IC_{50} values below 10 μM, although these compounds did not stabilize the DRH sequence in the FRET melting assay and there was no other evidence for ligand-DNA interaction in both the MD simulations and CD studies. Further investigations would be required to obtain information on the mode of action of these compounds.

3B.4.2 MTT Assay Results with HeLa Cell Line (Cervical Cancer Cell Line)

As the confluency rate of HeLa cells is higher than those of the other cell lines studied, HeLa was seeded at a lower density (*i.e.*, 5,000 cells per well) for this experiment. All sixteen compounds were evaluated against this cell line, where six compounds showed IC_{50} values ranges from 1.2-8.8 μM (**Table 3B.8)**.

Table 3B.8: Cytotoxic activities of novel compounds from library-1, library-3B and library-4 in the HeLa cell line.

Library	Compound	IC_{50} Value (μM) * Mean (n=3)	R^2
Library-1	**3.4**	> 100 μM	0.5789
	3.5	75.0 μM	0.8320
	3.6	4.2 μM	0.8469
	3.7	> 100 μM	0.6079
	3.8	2.7 μM	0.5789
Library-3B	**3.39**	28.6 μM	0.9311
	3.42	8.5 μM	0.9413
	3.43	1.2 μM	0.9429
Library-4	**3.66**	> 100 μM	0.7295
	3.68	> 100 μM	0.9524
	3.69	> 100 μM	0.5601
	3.71	3.9 μM	0.8554
	3.73	8.8 μM	0.4796
	3.75	21.5 μM	0.7190
	3.77	> 100 μM	0.6579
	3.78	> 100 μM	0.5343

For the HeLa cells, the *bis*-benzimidazole molecule **3.8** (library-1) had the best antiproliferative activity (*i.e.*, IC_{50} = 2.7 μM). Compound **3.6** also had potent antiproliferative activity with an IC_{50} value of 4.2 μM. In addition, **3.43** gave an IC_{50} value of 1.2 μM, although this did not correlate with the molecular modelling and dynamics results. On the other hand, among the other selected compound from the same library, the *bis*-benzofuran-type molecule **3.71** (library-4) gave the greatest inhibition of HeLa cell proliferation with an IC_{50} value of 3.9 μM.

3B.4.3 MTT Assay Results with NCI H1975 Cell Line (Non-Small Lung Cancer Cell Line)

The cell density of NCI H1975 seeding for the MTT assay was 10,000 cells/well. Less than 50% of the compounds evaluated showed antiproliferative activity toward this cell line with library-1 and library-3B compounds showing greater activity than those of library-4. Among the tested compounds, **3.8** showed most promising activity with an IC_{50} value of 1.5 µM.

Table 3B.9: Cytotoxic activity of novel compounds in library-1, library-3B and library-4 against the NCI H1975 cell line.

Library	Compound	IC_{50} Value (µM) * Mean (n=3)	R^2
Library-1	**3.4**	> 100 µM	0.8420
	3.5	16.8 µM	0.9623
	3.6	11.8 µM	0.9623
	3.7	15.3 µM	0.9398
	3.8	1.5 µM	0.9343
Library-3B	**3.39**	49.3 µM	0.8490
	3.42	> 100 µM	0.9420
	3.43	5.1 µM	0.9006
Library-4	**3.66**	> 100 µM	0.5931
	3.68	> 100 µM	0.5778
	3.69	> 100 µM	0.6982
	3.71	> 100 µM	0.7663
	3.73	> 100 µM	0.6796
	3.75	75.0 µM	0.6732
	3.77	> 100 µM	0.6795
	3.78	> 100 µM	0.8393

The results from across the three cell lines indicate that compounds **3.6**, **3.8** and **3.43** are highly cytotoxic molecules, and the viability of cultured MDA-MB-231, HeLa and NCI H1975

cells was significantly decreased during 72 h incubation in the presence of these compounds (IC_{50} values ranged from 0.5 to 11.8 μM). The data suggested that **3.8** was the most cytotoxic candidate across the three cell lines with the lowest range of IC_{50} values (*i.e.*, 1.5 μM to 4.2 μM). The variation in structure of these molecules could be related to their activity. For example, the methylene spacers of compounds **3.4**, **3.6** and **3.8** are 6-, 8- and 10-methylenes, respectively (**Table 3A.1**). According to the IC_{50} values for these three compounds, those with longer methylene spacers had the greater cytotoxic activities and this also correlated with the FRET melting assay, and the MD and CD data. A study by Wright and co-workers demonstrated higher relative telomerase activity and hTERT mRNA in the MDA-MB-231 cell line compared with HeLa cell line.[35] Therefore, the greater activity in the MDA-MB-231 breast cancer cells could be due to the ability of these novel compounds to inhibit telomerase.

In addition, data analyzed using GraphPad Prism software showed a good fit of the data quantified in terms of R^2 and the precision of fit value in terms of 95% confidence interval.[339] The low R^2 and wide 95% confidence interval indicated low variation in the data. Even though there was no general rule for R^2, its expected value in this study was more than 0.7. The possible reason for this was the variation of cell density in each well.

3B.5 Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) Assay

The reverse transcriptase-PCR assay was used to determine the effect of the 'hit' compounds on the expression of genes relating to the upregulation and/or downregulation of telomerase. The genes of interest in this analysis were hTERT and cyclin B.

It has already been demonstrated that there is a relationship between apoptosis and telomerase activity, and several anticancer agents induce apoptosis by downregulating telomerase activity. Sebastian and co-workers also reported that telomerase activation is an important step in human carcinogenesis.[293] In contrast, suppression of this enzyme's activity has a key role in tumour suppression in normal human somatic cells.[157] Therefore, abnormally active telomerase has become an attractive target for cancer treatment. Additionally, downregulation of either hTERT[340] or hTR[341] expression has been reported as associated with the induction of apoptosis. hTERT, the telomeric reverse transcriptase, is the catalytic subunit

of telomerase.[35, 342] hTERT acts as an especially important factor in upregulating telomerase activity in human cancer cells. Wu and co-workers have reported that the hTERT promoter is rich in c-Myc binding sites which can mediate hTERT transcriptional activation directly.[155] This finding is also consistent with another report suggesting a link between c-Myc overexpression and induction of telomerase activity.[343]

Moreover, Kang and co-workers have reported that the expression of the hTERT subunit comprising two putative Akt-phosphorylation sites is also upregulated in several cancer cells, where Akt-kinase has been identified as an enhancer of telomerase activity through the phosphorylation of hTERT.[344] In a study by Sun and co-workers, the authors established the dose-dependent downregulation of the hTERT gene and telomerase activity in cisplatin-treated human ovarian cancer cells.[312] There are a number of studies showing that expression of hTERT is associated with the level of telomerase activity.[245, 312, 313, 345-347] For these reasons, it was decided to study the effect of a number of selected compounds on hTERT expression and telomerase activity.

Cyclin B is a mitotic cyclin, a class of protein that controls cell progression during the cell cycle by initiating cyclin-dependent kinase (Cdk) enzymes.[348] During the cell cycle, the amount of cyclin B and the activity of the cyclin B-Cdk complex (also known as maturation promoting factor or mitosis promoting factor, MPF) increases until the point of mitosis, when levels drop sharply due to degradation of the cyclin B-Cdk complex.[349] Therefore, cyclin B plays an essential role in a variety of cancer types,[350, 351] as it is essential for cells to enter mitosis and, consequently, for cell division. Thus, the level of cyclin B can deregulate cell growth leading to the formation of tumours. In contrast, if cyclin B levels drop, then the MPF complex cannot form, and cells cannot enter the M-phase which leads to a slowing of cell division.[319,320]

In addition, Begnami and co-workers have reported that the high levels of cyclin B in gastric cancers are indicative of poor diagnosis and lymph node metastasis.[352]

In 2011, McKerlie and co-workers reported that Cdk1 specifically phosphorylates the threonine 371 (T371) residue of TRF1 (telomeric repeat binding factor 1) during the mitotic cell cycle, and that TRF1 is involved in the regulation of the telomere structure. Such phosphorylation is

related to the loss of telomere-bound TRF1 and TIN2 proteins, consequently causing a reduction in telomere heterochromatin.[353] Thus cyclin B can be used as a surrogate biomarker for telomerase expression within cancer cells. Therefore, cyclin B expression was studied in order to understand the mechanism of action of the tested compounds on tumour cell division. The most active four compounds (**3.6**, **3.8**, **3.43** and **3.71**) from the cytotoxicity assay were studied for their effects on gene expression in tumour cells using a RT-PCR protocol.

3B.5.1 hTERT Expression

The qualitative analysis of gene expression was carried out by observing a decrease in band intensity in the sample lane (D+) compared with the LPS control (L+). Suppression of the hTERT gene was detected for compounds **3.6** (at 1 μM), **3.8** (at 1 μM and 5 μM) and **3.71** (at 10 μM) (**Figure 3B.29 a, b, c** and **e**). The band intensity for **3.8** (at 1 μM) was a little more intense, however at 5 μM, hTERT expression was completely inhibited. On the other hand, compound **3.43** was inactive, and no suppression of hTERT expression was obtained (**Figure 3B.29 c**).

In 2001, Damm and colleagues reported lists of genes which were downregulated and upregulated by long-term treatment (*i.e.*, 7, 28 and 56 days) with a telomerase inhibitor.[245] Their study revealed that an effective telomerase inhibitor can downregulate hTERT expression up to 5-fold. In our study, cells were exposed to the tested compounds for 8 h, and during that short time of exposure compounds **3.6** and **3.8** were observed to have some inhibitory effect on hTERT expression. Moreover, the exposure period was less than the incubation time in this study, a similar protocol to that followed in the cytotoxicity evaluation experiment.

In summary, based on these preliminary results, it can be concluded that molecules **3.6** and **3.8** are promising candidates for further evaluation for telomerase inhibitory activity.

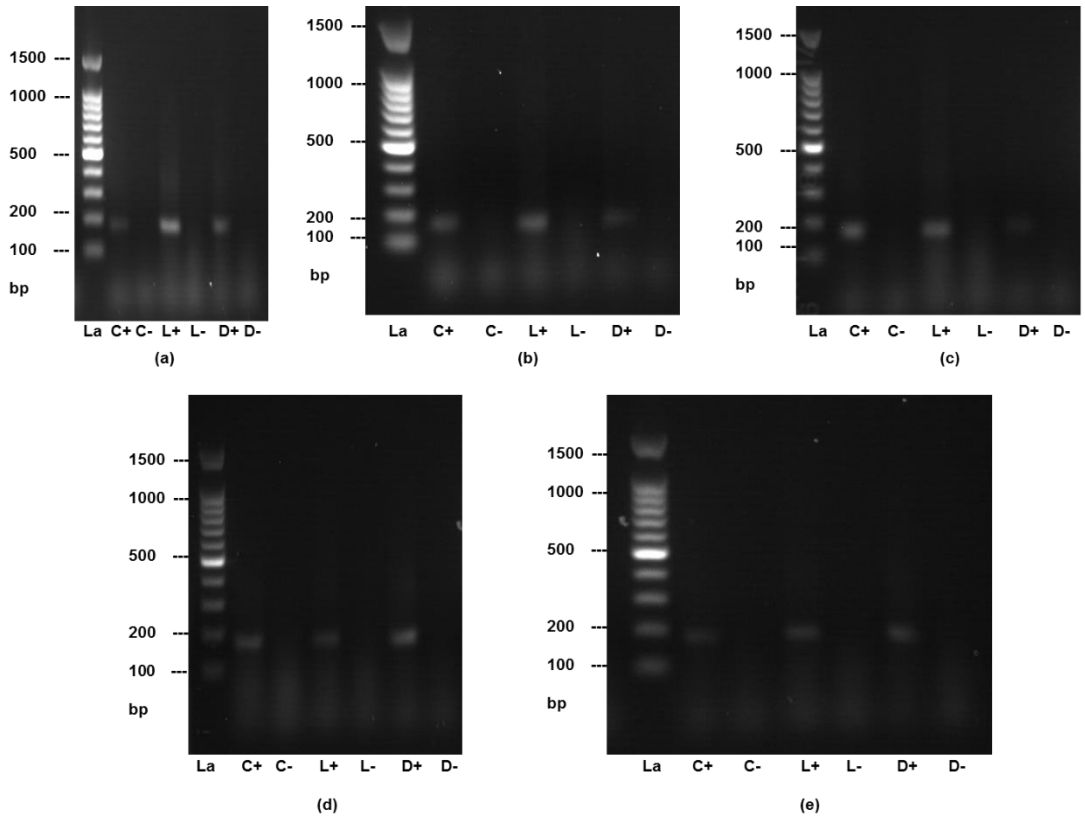


Figure 3B.29: Reversed Transcriptase-PCR results for modulation of hTERT gene expression by (a) Compound **3.6** (1 μM), (b) Compound **3.8** (1 μM), (c) Compound **3.8** (5 μM), (d) Compound **3.43** (1 μM) and (e) Compound **3.71** (10 μM). La: 100 bp DNA ladder, C : cell control, L: Cell stimulated with LPS, D: Cell stimulated with LPS and compound, + : RT enzyme, -: No RT enzyme.

3B.5.2 Cyclin B Expression

Similarly, four compounds, **3.6**, **3.8**, **3.43** and **3.71**, were evaluated using the same RT-PCR assay for effects on cyclin B expression in MDA-MB-231 cells. The gels showed a slight downregulation of cyclin B upon addition of compound **3.6** at a 1 μM concentration (**Figure 3B.30 a**). However, the other ligands, **3.8**, **3.43** and **3.71**, did not cause any significant changes in terms of band intensity at a 1 μM concentration. For this reason, compounds **3.8** and **3.71** were assessed at the higher concentrations of 5 μM and 10 μM, respectively. At a 5 μM concentration, there was a very small change in the intensity of the band for compound **3.8** (**Figure 3B.30 c**).

It is noteworthy that Damm and co-workers reported downregulation of cyclin B in NCI-H460 cells after prolonged treatment (*i.e.*, 7, 28 and 56 days) with the telomerase inhibitor BIBR1591.[245] In the future, longer incubation times for compounds **3.8**, **3.43** and **3.71** could be investigated.

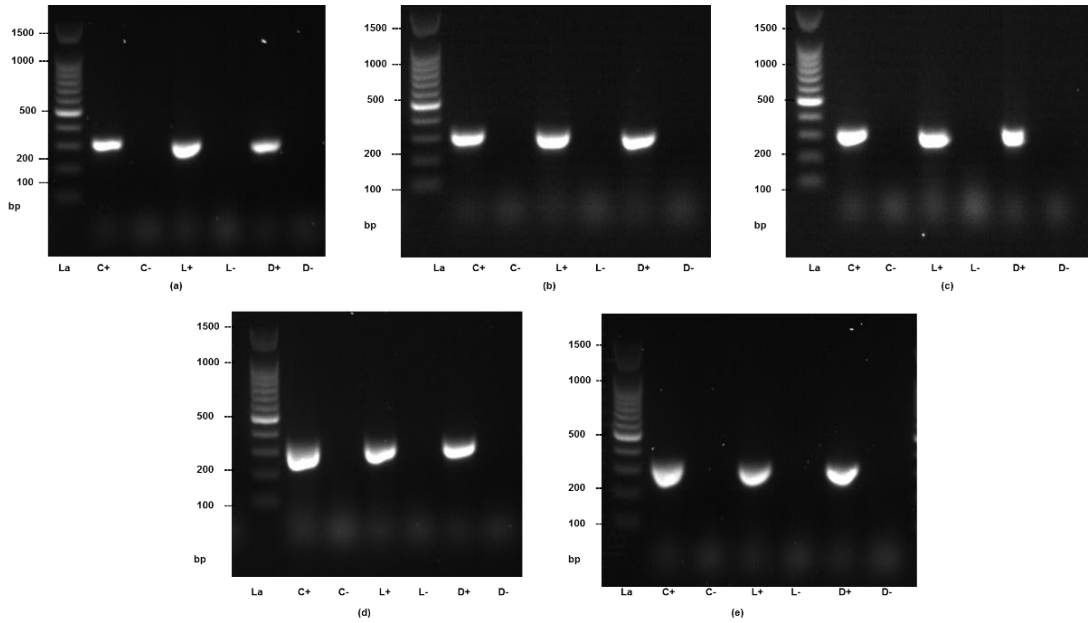


Figure 3B.30: RT-PCR results for evaluation of the effect on Cyclin B gene expression in MDA-MB-231 cells by (a) Compound **3.6** (1 μM), (b) Compound **3.8** (1 μM), (c) Compound **3.8** (5 μM), (d) Compound **3.43** (1 μM) and (e) Compound **3.71** (10 μM). La: 100 bp DNA ladder, C: cell control, L: Cell stimulated with LPS, D: Cell stimulated with LPS and compound, +: RT enzyme, -: No RT enzyme.

In 2008, Androic and co-workers demonstrated an association between cyclin B expression and the proliferation of cervical and breast cancer cells which were inhibited by the downregulation of cyclin B.[350] In the following year, Aaltonen and co-workers reported a study correlating cyclin B expression and breast cancer proliferation.[351] In 2010, Kim and co-workers reported that the highest level of cyclin B expression is observed in the G2/Mitosis transition.[354] As compound **3.6** downregulated the expression of cyclin B in MDA-MB-231 cells, it is possible that **3.6** may block cell cycle progression into the M-phase, thus causing G2 cell cycle arrest.

In summary, the preliminary gene expression experiments carried out in this project using the RT-PCR technique suggested that compound **3.6** could down-regulate the hTERT and cyclin B expression in treated cell lines.

3B.6 Electrophoretic Mobility Shift Assay

An electrophoretic mobility shift assay (EMSA, or gel shift assay) is an affinity electrophoresis method that is used to study protein/DNA and protein/RNA interactions. Although this assay is normally used to determine the capability of a protein or mixture of proteins to bind to a DNA or RNA sequence, the EMSA can also be applied to the study of the interaction of small molecules with DNA, RNA or DNA/RNA hybrid complexes. However, there are no previously published examples of its use to study the interaction of small molecules with DNA/RNA hybrid duplexes.

According to the electrophoretic mobility shift assay (EMSA) principle, it has been confirmed that the protein retards the mobility of the DNA fragments to which it binds; thus, the free DNA migrates faster through the gel than does the DNA-protein complex.[355] Similarly, this method has been used by different research group to study drug/DNA binding with the idea of DNA migration with/without drug.

In our study, the gel showed a very slight retardation of the band after treatment with **3.6** and **3.8** at a 2.5 μM ligand concentration compared with the control DNA ladder (without ligand) (**Figure 3B.31**), which suggests the binding of ligand to the treated DNA strand. This result recommends that future experiments could focus on studying higher concentrations of ligands to establish whether greater shifts can be obtained.

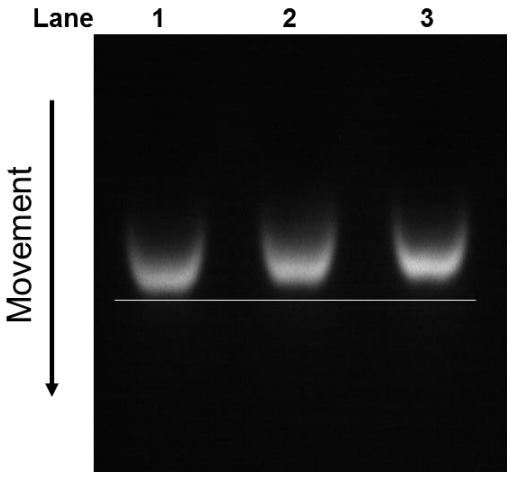


Figure 3B.31: Electrophoretic mobility shift assay using 4% agarose gel. Control Oligo-DRH (lane 1), Oligo-DRH with compound **3.6** (2.5 μM) (in lane 2), Oligo-DRH with compound **3.8** (2.5 μM) (in lane 3).

Chapter 4: Conclusion and Future Work

After careful modification of the NCI library molecule NSC 273829, a total of 49 molecules with varied lengths and flexibilities of methylene spacer and with seven types of terminal building blocks were synthesized using a general HOBt/DIC-mediated amide coupling reaction. Most of the target molecules were achieved in good yield and purity using established medicinal chemistry protocols. Several spectroscopic techniques such as ^{1}H NMR, ^{13}C NMR, IR, LC-MS and HRMS were employed to confirm the structures of the synthesized molecules. The initial lead molecule was modified with the help of molecular modelling studies to design the first library.

Compounds **3.2**-**3.10** (library-1) were synthesized by varying the linker length (methylene spacer) as well as by inserting a rigid-planar structure of phenyls and biphenyls in place of the flexible methylene spacer. Preliminary FRET melting assays on dual labelled telomeric DNA/RNA hybrid (DRH) duplex sequences, control DNA duplex (cDRH) and telomeric G-quadruplex sequences (F21T) were carried out to assess the ability of the library members to stabilize the DRH sequence and to understand their selectivity.

Compounds **3.4** (six methylene spacer), **3.6** (eight methylene spacer) and **3.8** (ten methylene spacer) showed promising selectivity, 11.5-, 9.0- and 7.6-fold, respectively toward the DRH sequence compared with the control duplex sequence (cDRH).

Further molecular dynamics studies also confirmed that the molecules with an even number of methylene spacers (for example, six, eight, ten) were more selective toward the DRH sequence compared with the compounds with an odd number of methylene spacers. It was also shown that in their lowest possible energy conformations, compounds **3.6** and **3.8** can stabilize the DRH sequence by a dual mode of action, *bis*-intercalation into the nucleobases and hybrophobic interactions with the minor groove of the DRH sequence, whereas compound **3.4** could mono-intercalate into the nucleobases, with the second benzimidazole pointing out from the DNA bases. In addition, MD studies confirmed that the single methylene spacer unit

between the phenylene and benzimidazole moieties is responsible for forming a kink and allows the benzimidazole moiety to intercalate.

CD spectroscopic analysis also showed notable shifting in the positive signal upon gradual addition of ligands (compounds **3.6** and **3.8**) compared with the CD signal for the DRH itself. There is no published CD study on the telomeric DRH sequence and therefore it was not possible to compare the observed changes in CD with other literature molecules.

Synthesized molecules were assayed for their cytotoxic potential against the breast cancer cell line MDA-MB-231. On the basis of the preliminary investigation, sixteen promising molecules from libraries -1, -3B and -4 were further evaluated against the cervical cancer cell line HeLa and non-small lung cancer cell line NCI H1975, in which compounds **3.6** and **3.8** gave low micromolar IC_{50} values for all three cell lines used.

In addition, RT-PCR assays for hTERT and cyclin B expression were performed with compounds **3.6** and **3.8**, where both compounds down-regulated hTERT expression at 1 μM concentration. However, compound **3.8** did not suppress cyclin B expression like compound **3.6** at 1 μM concentration. An electrophoretic mobility shift assay (EMSA) was also performed to evaluate the ability of compounds **3.6** and **3.8** to bind to the DRH sequence.

The result suggested that the compounds could bind to the DRH sequence and thus showed differences in distance travelled by the ligand-bound DNA compared with the DRH itself. Though a well-established TRAP assay for telomerase inhibition could not be included in this thesis due to time constrains, it should be noted that promising compounds will be evaluated in the TRAP-LIG assay in the future. Having obtained evidence of ligand-DRH interaction at a low micromolar level, it can be hypothesised that such molecules can be used as leads for further investigation in terms of binding selectivity, cytotoxicity and drug-like properties which could lead to potential therapeutic agents in the oncology area.

As some of the compounds synthesized have demonstrated significant activity against the breast cancer cell line MDA-MB-231, further biological investigations will be carried out to establish the mechanism of action. Molecular dynamics studies suggest that the single methylene spacer between the phenylene and benzimidazole moieties is responsible for *bis*-

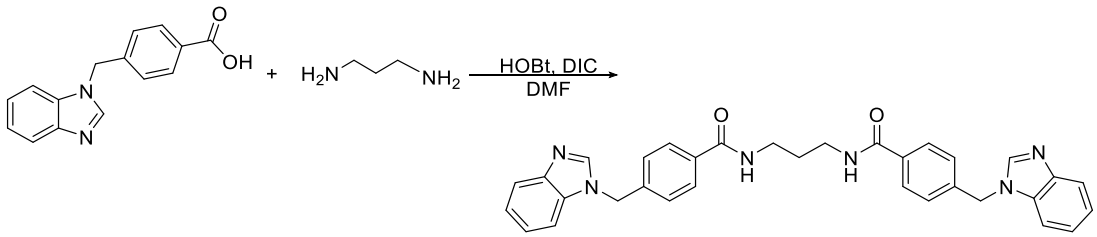
intercalation, further SAR studies will involve varying and/or inserting additional methylene spacers, as well as utilizing different terminal building blocks. Future work will also focus on establishing a detailed mechanism of action of these ligands at the cellular level so that the 'lead' compounds can undergo further modification to generate suitable preclinical candidates.

Appendix 1: Experimental, Synthesis and Characterisation of Molecules

A1.1 Synthesis of *bis*-benzo[*d*]imidazole type molecules in library-1

Synthesis of *N,N′*-(propane-1,3-diyl)bis(4-((1*H*-benzo[*d*]imidazol-1-l)methyl)benzamide) (3.2)

A solution of 4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzoic acid (HOBt) (50 mg; 0.20 mmol) in dimethylformamide (7 mL) was treated with 1-hydroxybenzotriazole (54 mg, 0.40 mmol) and *N,N′*-diisopropylcarbodiimide (DIC) (54 μL, 0.35 mmol) at room temperature. After 30 minutes, 1,7-diaminopropane (31 mg, 0.24 mmol) was added and the mixture was stirred for 16 h, then quenched with water (70 mL). The product was extracted with ethyl acetate (3x100 mL) and the combined organic layer was washed with a saturated aqueous solution of sodium bicarbonate (2x150 mL), then dried with magnesium sulphate, filtered and concentrated under pressure to give the desired compound **3.2**.

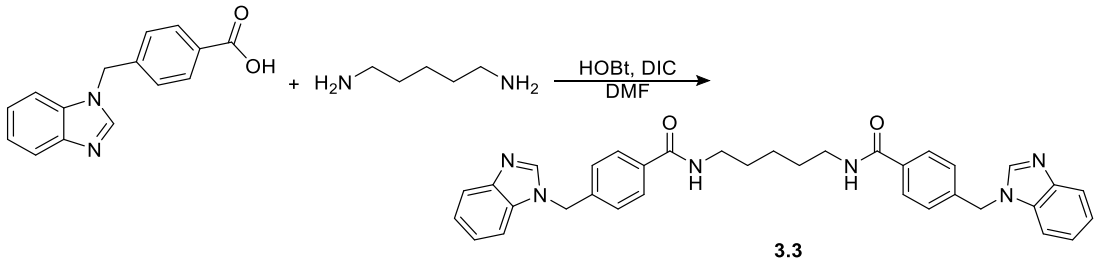


3.2

3.2 White solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.46-8.42 (m, 4 H), 7.79-7.75 (m, 4 H), 7.68-7.64 (m, 2 H), 7.49-7.45 (m, 2 H), 7.35 (d, *J* = 6.6 Hz, 4 H), 7.22-7.17 (m, 4 H), 5.55 (s, 4 H), 3.29-3.22 (m, 4 H), 1.70 (t, *J* = 6.3 Hz, 2 H),
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 166.1, 144.5, 143.8, 142.0, 141.6, 140.1, 134.3, 133.8, 127.7, 122.7, 119.7, 110.9, 47.5, 37.2, 29.4
	IR	(υ_{max}/cm^{-1}) 3344, 3074, 1647, 1616, 1570, 1543, 1496, 1377, 1365
	MS	Found 543.20 $[M+H]^+$, Calculated for $C_{33}H_{30}N_6O_2$ 543.64 $[M+H]^+$
	HRMS	Observed 543.2509 $[M+H]^+$, Theoretical 543.2503 $[C_{33}H_{30}N_6O_2+H]^+$
	Yield	52 mg (29% yield)

Synthesis of *N,N′*-(pentane-1,5-diyl)bis(4-((1*H*-benzo[*d*]imidazol-1-l)methyl)benzamide) (3.3)

Protocol employed for the preparation of **3.2** was used. Purification was carried out in flash column chromatography using dichloromethane/methanol (0-3%).

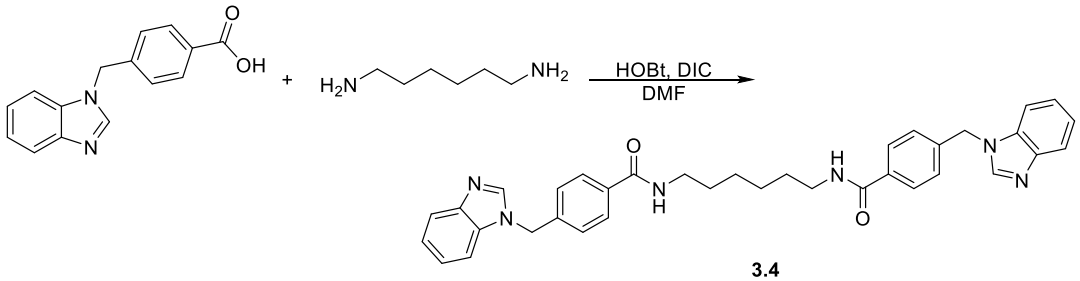


3.3

3.3 White solid	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 8.09 (s, 2 H), 7.79 (d, *J* = 7.3 Hz, 2 H), 7.83 (d, *J* = 8.6 Hz, 2 H), 7.72 (d, *J* = 8.4 Hz, 4 H), 7.36-7.30 (m, 4 H), 7.12 (d, *J* = 8.6 Hz, 4 H), 6.50 (t, *J* = 6.2 Hz, 2 H), 5.38 (s, 4 H), 3.47 (q, *J* = 6.2 Hz, 4 H), 1.71-1.65 (m, 4 H), 1.48-1.43 (m, 2 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 163.0, 145.4, 143.8, 141.3, 133.9, 130.1 (2 C), 127.9 (2 C), 127.2, 124.3, 123.8, 120.6, 110.6, 50.2, 39.1, 31.3, 24.5
	IR	(υ_{max}/cm^{-1}) 3268, 2936, 1638, 1616, 1540, 1504, 1456, 1385, 1364
	MS	Found 570.83 $[M-H]^+$, Calculated for $C_{35}H_{34}N_6O_2$ 571.68 $[M+H]^+$
	HRMS	Observed 571.2824 $[M+H]^+$, Theoretical 571.2816 $[C_{35}H_{34}N_6O_2+H]^+$
	Yield	26 mg (23% yield)

Synthesis of *N,N′*-(hexane-1,6-diyl)bis(4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzamide) (3.4)

Compound **3.4** was obtained by following procedure used in the synthesis of **3.2**.

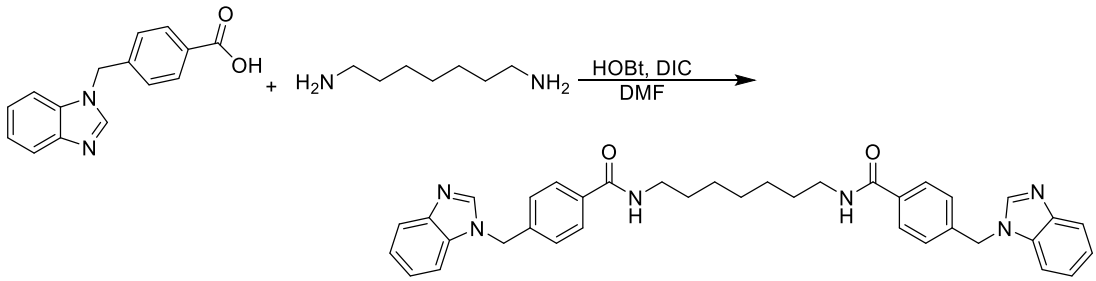


3.4

3.4 White solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.42 (s, 2 H), 8.37 (t, *J* = 5.7 Hz, 2 H), 7.76 (d, *J* = 8.3 Hz, 4 H), 7.68-7.64 (m, 4 H), 7.35 (d, *J* = 8.3 Hz, 4 H), 7.20-7.17 (m, 4 H), 5.55 (s, 4 H), 3.22-3.17 (m, 4 H), 1.46 (br s, 4 H), 1.28 (br s, 4 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 166.6, 144.4, 143.2, 141.7, 133.0, 130.2 (2 C), 127.5 (2 C), 127.2, 122.6, 121.9, 120.4, 110.6, 59.8, 39.7, 29.6, 26.8
	IR	(υ_{max}/cm^{-1}) 3285, 2935, 2359, 1641, 1569, 1540, 1456, 1436, 1363
	MS	Found 584.50 $[M-H]^+$, Calculated for $C_{36}H_{36}N_6O_2$ 585.71 $[M+H]^+$
	HRMS	Observed 585.2964 $[M+H]^+$, Theoretical 585.2973 $[C_{36}H_{36}N_6O_2+H]^+$
	Yield	65 mg (52% yield)

Synthesis of *N,N'*-(heptane-1,7-diyl)bis(4-((1*H*-benzo[*d*]imidazol-1-l)methyl)benzamide) (3.5)

Product **3.5** was prepared by following the same protocol used in the synthesis of compound **3.2**. Purification was carried out by crystallization with water to give the desired compound as white solid.

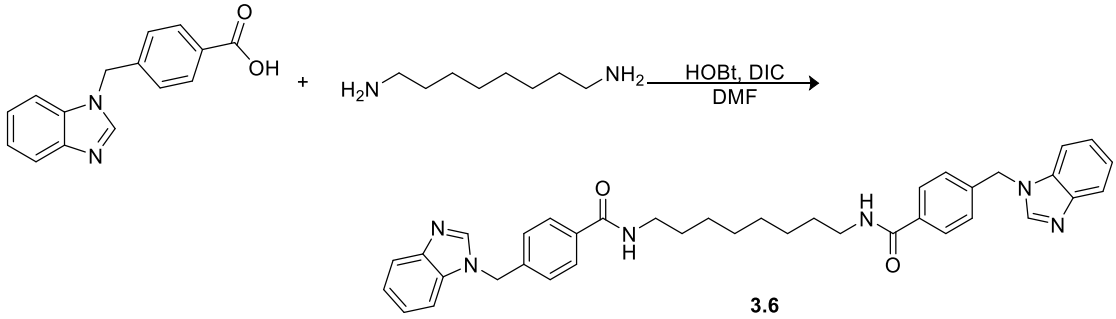


3.5

3.5 Pale yellow solid	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 7.95 (s, 2 H), 7.82-7.78 (m, 2 H), 7.75-7.70 (m, 4 H), 7.23-7.21 (m, 8 H), 7.17 (d, *J* = 8.0 Hz, 4 H), 4.20 (br s, 4 H), 3.86-3.80 (m, 2 H), 3.43-3.36 (m, 2 H), 1.56 (t, *J* = 6.5 Hz, 4 H), 1.34 (br s, 4 H), 1.21 (d, *J* = 6.5 Hz, 2 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 166.5, 143.5, 142.8, 138.4, 134.5, 133.4, 127.3 (2 C), 126.7 (2 C), 122.9, 122.1, 120.1, 109.6, 48.1, 30.5, 28.9, 28.2, 26.2
	IR	(υ_{max}/cm^{-1}) 3339, 2966, 1614, 1556, 1494, 1458, 1361
	MS	Found 599.03 $[M+H]^+$, Calculated for $C_{37}H_{38}N_6O_2$ 599.74 $[M+H]^+$
	HRMS	Observed 599.3123 $[M+H]^+$; Theoretical 599.3129 $[C_{37}H_{38}N_6O_2+H]^+$
	Yield	99 mg (83% yield)

Synthesis of *N,N'*-(octane-1,8-diyl)bis(4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzamide) (3.6)

Compound **3.6** was obtained as a white solid after 18 h of reaction following the procedure employed for the synthesis of compound **3.2**.



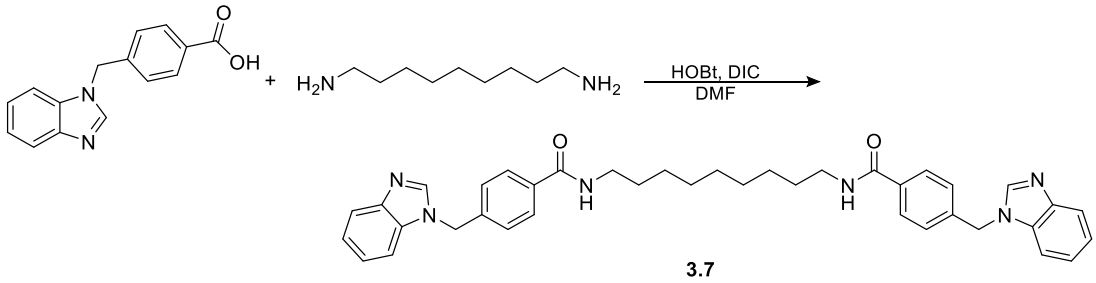
3.6

3.6 White solid	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 7.98 (s, 1 H) , 7.84 (d, *J* = 7.8 Hz, 1 H), 7.72 (d, *J* = 8.4 Hz, 2 H), 7.28 (d, *J* = 2.5 Hz, 1 H), 7.26-7.23 (m, 2 H), 7.21 (d, *J* = 8.4 Hz, 2 H), 6.13 (t, *J* = 5.9 Hz, 1 H), 5.41 (s, 2 H), 4.00 (br s, 2 H), 3.84 (dd, *J* = 13.1, 6.6 Hz, 5 H), 3.45-3.39 (m, 2 H), 1.62-1.56 (m, 2 H), 1.34 (br s, 2 H), 1.16 (m, 15 H)

	^{13}C NMR	(100 MHz, $CDCl_3$) δ 161.0, 148.1, 143.3, 142.8, 138.4, 134.2, 127.3 (2 C), 126.9 (2 C), 123.1, 122.3, 120.2, 109.6, 48.2, 39.7, 29.2, 28.6, 26.4
	IR	(υ_{max}/cm^{-1}) 3328, 2928, 1627, 1562, 1496, 1458, 1383, 1325
	MS	Found 613.88 $[M+H]^+$, Calculated for $C_{38}H_{40}N_6O_2$ 613.76 $[M+H]^+$
	HRMS	Observed 613.3268 $[M+H]^+$, Theoretical 613.3286 $[C_{38}H_{40}N_6O_2+H]^+$
	Yield	78 mg (74% yield)

Synthesis of *N,N'*-(nonane-1,9-diyl)bis(4-((1*H*-benzo[*d*]imidazol-1-yl)methyl)benzamide) (3.7)

General method described for the synthesis of compound **3.2** was used. To ensure the completion of reaction, additional DIC (26.8 mL, 0.21 mmol) and HOBt (27 mg, 0.2 mmol) were needed.

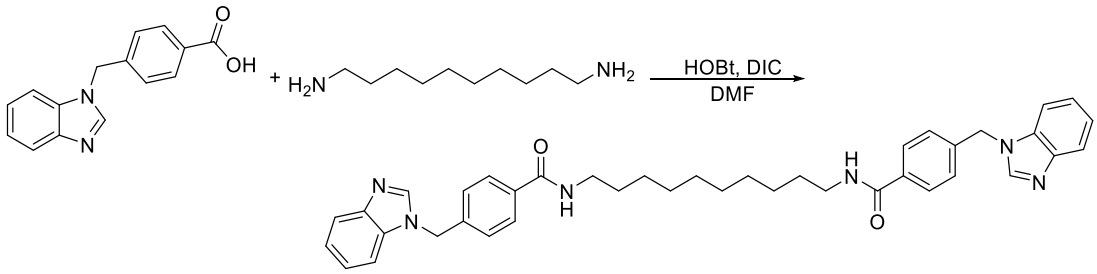


3.7

3.7 Pale yellow solid	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 7.97 (s, 2 H), 7.84 (d, *J* = 8.1 Hz, 4 H) 7.72 (d, *J* = 8.4 Hz, 4 H) 7.24 (td, *J* = 2.9, 1.1 Hz, 4 H) 7.21 (d, *J* = 8.1 Hz, 4 H) 6.20-6.15 (m, 2 H) 5.41 (s, 4 H) 3.87-3.83 (m, 4 H) 3.45-3.39 (m, 4 H) 1.27 (br s, 10 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 162.2, 143.9, 142.8, 142.0, 134.1, 127.6, 127.3, 126.8, 125.5, 122.9, 120.2, 109.6, 76.9, 50.5, 36.1, 31.1, 28.6, 23.2
	IR	(υ_{max}/cm^{-1}) 3335, 2966, 2360, 2338, 1638, 1556, 1495, 1458, 1384, 1325
	MS	Found 626.98 $[M+H]^+$, Calculated for $C_{39}H_{42}N_6O_2$ 626.79 $[M+H]^+$
	HRMS	Observed 627.3436 $[M+H]^+$, Theoretical 627.3442 $[C_{39}H_{42}N_6O_2+H]^+$
	Yield	32 mg (31% yield)

Synthesis of *N,N'*-(decane-1,10-diyl)bis(4-((1*H*-benzo*[*d]imidazol-1-l)methyl)benzamide) (3.8)

General method described for the synthesis of compound **3.2** was used. Product was collected by filtration when ethyl acetate was added.

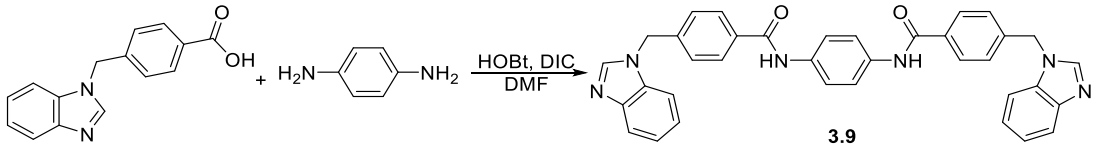


3.8

3.8 White solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.06 (s, 2 H), 7.86 (d, *J* = 8.0 Hz, 4 H), 7.68 (d, *J* = 8.3 Hz, 4 H), 7.22 (td, *J* = 2.9, 1.1 Hz, 4 H), 7.20 (d, *J* = 8.0 Hz, 4 H), 6.20-6.15 (m, 2 H), 5.41 (s, 4 H), 3.87-3.83 (m, 4 H), 3.45-3.39 (m, 4 H), 1.26 (s, 12 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 166.5, 144.2, 143.4, 141.8, 135.1, 131.2, 128.6, 127.6, 122.6, 121.9, 120.4, 110.6, 51.5, 41.0, 31.1, 29.2, 28.6, 26.2
	IR	(υ_{max}/cm^{-1}) 3336, 2924, 2850, 1639, 1616, 1543, 1492, 1438, 1369
	MS	Found 641.30 $[M-H]^+$, Calculated for $C_{40}H_{44}N_6O_2$ 641.35 $[M+H]^+$
	HRMS	Observed 641.3580 $[M+H]^+$, Theoretical 641.3599 $[C_{40}H_{44}N_6O_2+H]^+$
	Yield	78 mg (70% yield)

Synthesis of *N,N'*-(1,4-phenylene)bis(4-((*1H*-benzo[*d*]imidazol-1-yl)methyl)benzamide) (3.9)

Compound **3.9** was synthesized by following general amide coupling reaction used in **3.2**. Purification was done by using similar protocol used for the compound **3.2**.

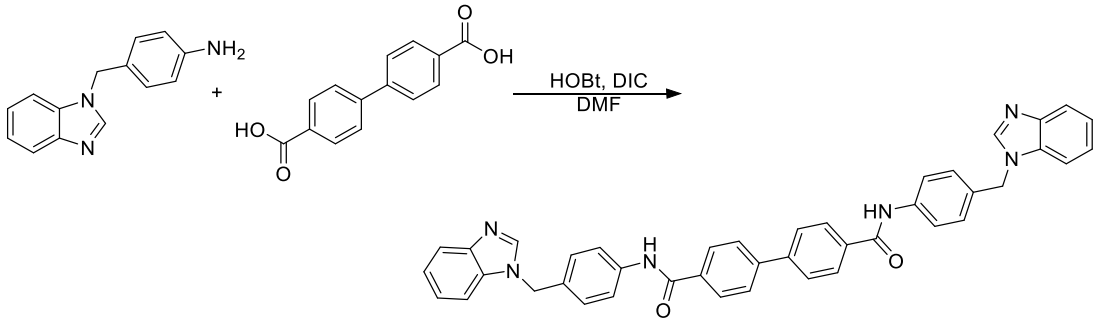


3.9

3.9 White powder	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 10.19 (s, 2 H), 8.46 (s, 2 H), 7.88 (d, *J* = 8.1 Hz, 4 H), 7.71-7.64 (m, 6 H), 7.51 (d, *J* = 8.6 Hz, 2 H), 7.43 (d, *J* = 8.3 Hz, 4 H), 7.23-7.17 (m, 4 H), 5.60 (s, 4 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 164.9, 143.9, 140.4, 134.9, 134.5, 131.8, 128.1 (2 C), 127.4 (2 C), 123.7 (2 C), 122.5, 121.7, 120.5, 110.7, 59.9
	IR	(υ_{max}/cm^{-1}) 3286, 3163, 3066, 2947, 1612, 1570, 1492, 1458, 1350
	MS	Found 577.00 $[M+H]^+$, Calculated for $C_{36}H_{28}N_6O_2$ 577.66 $[M+H]^+$
	HRMS	Observed 577.2371 $[M+H]^+$; Theoretical 577.2347 $[C_{36}H_{28}N_6O_2+H]^+$
	Yield	81 mg (51% yield)

Synthesis of *N*$_4$,*N*$_4$'-bis(4-((1H-benzo[*d*]imidazol-1-yl)methyl)phenyl)-[1,1'-biphenyl]-4,4'-dicarboxamide (3.10)

Protocol used for **3.9** was followed.



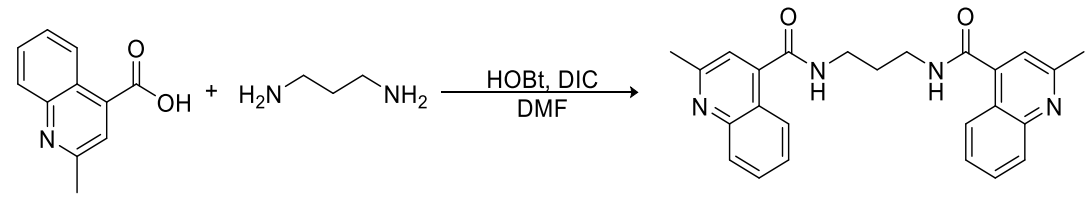
3.10

3.10 White solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 10.47 (s, 2 H), 9.69 (s, 2 H), 8.11-8.07 (m, 4 H), 7.94-7.91 (m, 4 H), 7.90-7.86 (m, 4 H), 7.85-7.82 (m, 4 H), 7.58-7.54 (m, 4 H), 7.51 (d, *J* = 8.6 Hz, 4 H), 5.72 (s, 4 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 164.2, 144.9, 142.2, 139.6, 134.9, 134.3, 133.1, 132.9, 128.9 (2 C), 128.7 (2 C), 127.1 (2 C), 125.9, 122.6, 122.0, 120.8 (2 C), 113.4, 51.5
	IR	(υ_{max}/cm^{-1}) 3105, 3032, 2989, 1597, 1493, 1438, 1415, 1385, 1365
	MS	Found 653.10 $[M+H]^+$, Calculated for $C_{42}H_{32}N_6O_2$ 653.26 $[M+H]^+$
	HRMS	Observed 653.2670 $[M+H]^+$; Theoretical 653.2660 $[C_{42}H_{32}N_6O_2+H]^+$
	Yield	65 mg (90% yield)

A1.2 Synthesis of *bis*-quinoline type molecules in library-2A

Synthesis of *N,N'*-(propane-1,3-diyl)*bis*(2-methylquinoline-4-carboxamide) (3.12)

To a solution of 2-methylquinoline-4-carboxylic acid (82 mg, 0.44 mmol) in dimethylformamide (7 mL), HOBt (112 mg, 0.83 mmol) and DIC (113 μL, 0.73 mmol) were successively added at room temperature and the mixture was allowed to stir for 30 minutes. 1,3-Diaminopropane (30 mg, 0.21 mmol) was added. After 6 h of stirring, reaction was quenched with water (10 mL) and extracted with ethyl acetate (3x20 mL). Organics were combined and washed with a saturated aqueous solution of sodium bicarbonate (20 mL), dried over magnesium sulphate, filtered and concentrated. The desired compound was purified by flash column chromatography using ethyl acetate/methanol (0-10%).

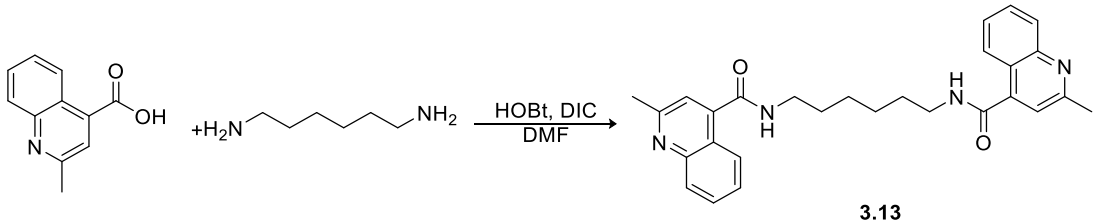


3.12

3.12 Amber solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.78 (t, *J* = 5.5 Hz, 2 H), 8.18 (dd, *J* = 8.1, 0.9 Hz, 2 H), 7.96 (t, *J* = 8.1 Hz, 2 H), 7.87 (ddd, *J* = 8.4, 7.0, 1.3 Hz, 2 H), 7.76 (ddd, *J* = 8.4, 7.0, 1.3 Hz, 2 H), 7.69 (s, 2 H), 3.15 (t, *J* = 6.1 Hz, 4 H), 2.61 (s, 6 H), 1.78-1.72 (m, 2 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 167.2, 157.9, 146.6, 142.4, 129.2, 128.8, 126.0, 125.4, 123.1, 120.1, 39.7, 28.8, 23.2
	IR	(υ_{max}/cm^{-1}) 3271, 3059, 2936, 1639, 1589, 1435, 1377, 1338
	MS	Found 413.00 $[M+H]^+$, Calculated for $C_{25}H_{24}N_4O_2$ 413.49 $[M+H]^+$
	HRMS	Observed 413.1973 $[M+H]^+$; Theoretical 413.1972 $[C_{25}H_{24}N_4O_2+H]^+$
	Yield	51 mg (48% yield)

Synthesis of *N,N′*-(hexane-1,6-diyl)*bis*(2-methylquinoline-4-carboxamide) (3.13)

Method described for the preparation of compound **3.12** was employed, desired product was collected by filtration when crude was treated with methanol and diethyl ether.

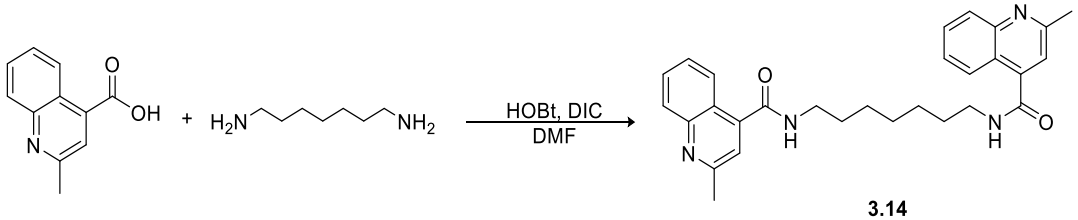


3.13

3.13 Tan powder	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.75 (s, 2 H), 8.07-8.03 (m, 2 H), 7.96 (d, *J* = 8.1 Hz, 2 H), 7.74 (ddd, *J* = 8.3, 7.1, 1.3 Hz, 2 H), 7.57 (ddd, *J* = 8.3, 7.1, 1.3 Hz, 2 H), 7.42 (s, 2 H), 3.34 (d, *J* = 4.0 Hz, 4 H), 2.67 (s, 6 H), 1.64-1.57 (m, 4 H), 1.43 (m, 4 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 166.7, 158.6, 147.6, 142.6, 129.7, 128.6, 126.3, 124.5, 122.6, 119.7, 35.2, 29.0, 26.2, 24.2
	IR	(υ_{max}/cm^{-1}) 3256, 3059, 2928, 1658, 1501, 1469, 1408, 1350
	MS	Found 455.10 $[M+H]^+$, Calculated for $C_{28}H_{30}N_4O_2$ 455.57 $[M+H]^+$
	HRMS	Observed 455.2441 $[M+H]^+$; Theoretical 455.2442 $[C_{28}H_{30}N_4O_2+H]^+$
	Yield	31 mg (26% yield)

Synthesis of *N,N′*-(heptane-1,7-diyl)*bis*(2-methylquinoline-4-carboxamide) (3.14)

Desired product **3.14** was obtained by using the protocol described in the synthesis of compound **3.12**. Purification was carried out by flash column chromatography with ethyl acetate/methanol (0-5%).

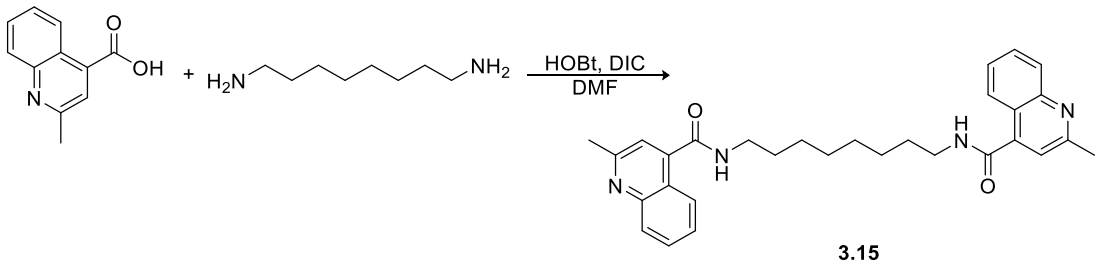


3.14

3.14 Tan solid	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 8.16 (dd, *J* = 8.4, 0.9 Hz, 2 H), 8.08 (d, *J* = 8.1 Hz, 2 H), 7.78 (ddd, *J* = 8.4, 6.9, 1.5 Hz, 2 H), 7.71 (s, 2H), 7.58 (ddd, *J* = 8.4, 6.9, 1.5 Hz, 2 H), 7.01 (t, *J* = 5.8 Hz, 2 H), 3.65-3.59 (m, 4 H), 2.69 (m, 6 H), 1.79 (t, *J* = 6.4 Hz, 4 H), 1.55 (m, 6 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 167.4, 158.3, 147.6, 142.6, 123.0, 128.4, 126.6, 124.9, 122.6, 119.3, 39.8, 29.3, 28.5, 26.6, 24.8
	IR	(υ_{max}/cm^{-1}) 3279, 3071, 2928, 1632, 1589, 1462, 1439, 1408, 1338
	MS	Found 468.90 $[M+H]^+$, Calculated for $C_{29}H_{32}N_4O_2$ 468.60 $[M+H]^+$
	HRMS	Observed 469.2603 $[M+H]^+$; Theoretical 469.2598 $[C_{29}H_{32}N_4O_2+H]^+$
	Yield	54 mg (50% yield)

Synthesis of *N,N*'-(octane-1,8-diyl)*bis*(2-methylquinoline-4-carboxamide) (3.15)

Method described for the synthesis of **3.12** was employed. The desired compound **3.15** was obtained after purification by flash column chromatography using dichloromethane/isopropyl alcohol (0-30%).

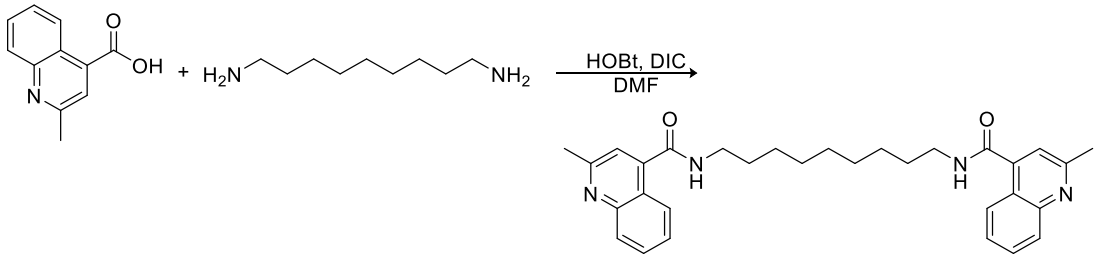


3.15

3.15 Tan solid	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 8.14-8.10 (m, 2 H), 8.03 (d, *J* = 8.6 Hz, 2 H), 7.71 (ddd, *J* = 8.4, 7.0, 1.5 Hz, 2 H), 7.56-7.50 (m, 2 H), 7.28 (s, 2 H), 6.18 (t, *J* = 5.5 Hz, 2 H), 3.58-3.52 (m, 4 H), 2.72 (s, 6 H), 1.73-1.68 (m, 4 H), 1.48-1.40 (m, 8 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 167.5, 158.5, 148.2, 142.4, 130.0, 129.0, 126.7, 124.9, 122.7, 119.3, 40.0, 29.5, 28.9, 26.8, 25.3
	IR	(υ_{max}/cm^{-1}) 3302, 2920, 2854, 1635, 1589, 1539, 1334
	MS	Found 483.8 $[M+H]^+$, Calculated for $C_{30}H_{34}N_4O_2$ 483.63 $[M+H]^+$
	HRMS	Observed 483.2741 $[M+H]^+$; Theoretical 483.2755 $[C_{30}H_{34}N_4O_2+H]^+$
	Yield	23.8 mg (24% yield)

Synthesis of *N,N'*-(nonane-1,9-diyl)*bis*(2-methylquinoline-4-carboxamide) (3.16)

The desired compound **3.16** was prepared following procedure described for the synthesis of **3.12** and obtained as a tan solid after purification by flash column chromatography using ethyl acetate/methanol/2M ammonia solution in methanol (80/15/5).

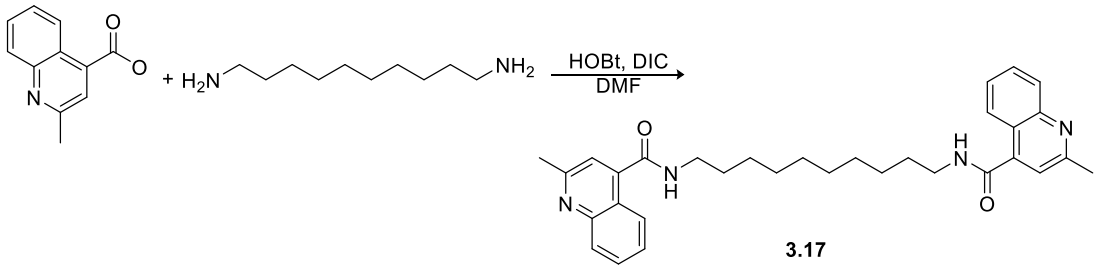


3.16

3.16 Pale yellow solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.78 (s, 2 H), 7.96 (t, *J* = 8.1 Hz, 4 H), 7.78 (ddd, *J* = 8.4, 6.9, 1.3 Hz, 2 H), 7.68 (ddd, *J* = 8.4, 6.9, 1.3 Hz, 2 H), 7.65 (s, 2 H), 3.14 (t, *J* = 4.1 Hz, 4 H), 2.58 (s, 6 H), 1.54-1.50 (m, 4 H), 1.27 (m, 10 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 165.2, 156.6, 148.2, 143.2, 130.9, 129.7, 126.4, 124.5, 123.6, 120.0, 37.3, 30.2, 29.4, 26.2, 23.4
	IR	(υ_{max}/cm^{-1}) 3290, 3059, 2924, 1635, 1589, 1466, 1435, 1377
	MS	Found 497.10 $[M+H]^+$, Calculated for $C_{31}H_{36}N_4O_2$ 497.66 $[M+H]^+$
	HRMS	Observed 497.2909 $[M+H]^+$; Theoretical 497.2911 $[C_{31}H_{36}N_4O_2+H]^+$
	Yield	80 mg (64% yield)

Synthesis of *N,N'*-(decane-1,10-diyl)*bis*(2-methylquinoline-4-carboxamide) (3.17)

Compound **3.17** was obtained as tan solid by following the general protocol used for **3.12**.



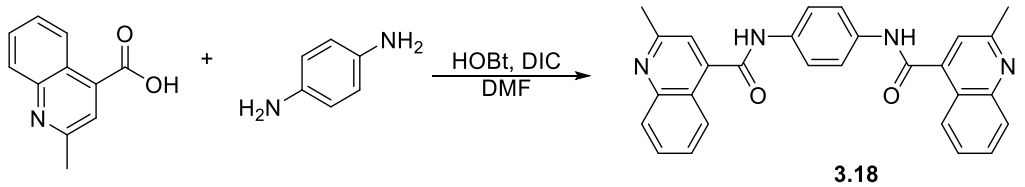
3.17

3.17 Tan solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.73 (t, *J* = 5.5 Hz, 2 H), 8.04 (dd, *J* = 8.2, 0.9 Hz, 2 H), 7.96 (d, *J* = 8.2 Hz, 2 H), 7.74 (ddd, *J* = 8.3, 6.9, 1.3 Hz, 2 H), 7.56 (ddd, *J* = 8.3, 6.9, 1.3 Hz, 2 H), 7.41 (s, 2 H), 3.32 (m, 4 H), 1.61-1.52 (m, 4 H), 2.67 (s, 6 H), 1.32 (m, 12 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 166.6, 158.5, 147.6, 142.6, 129.7, 128.6, 126.3, 125.2, 122.6, 119.6, 40.1, 29.0, 28.9, 28.8, 26.5, 24.8

	IR	(υ_{max}/cm^{-1}) 3309, 2924, 2854, 1635, 1585, 1473, 1442, 1323
	MS	Found 511.00 $[M+H]^+$, Calculated for $C_{32}H_{38}N_4O_2$ 511.68 $[M+H]^+$
	HRMS	Observed 511.3068 $[M+H]^+$; Theoretical 511.3068 $[C_{32}H_{38}N_4O_2+H]^+$
	Yield	49 mg (38% yield)

Synthesis of *N,N'*-(1,4-phenylene)*bis*(2-methylquinoline-4-carboxamide) (3.18)

Compound **3.18** was obtained as tan solid by following the protocol employed to prepare **3.12**.



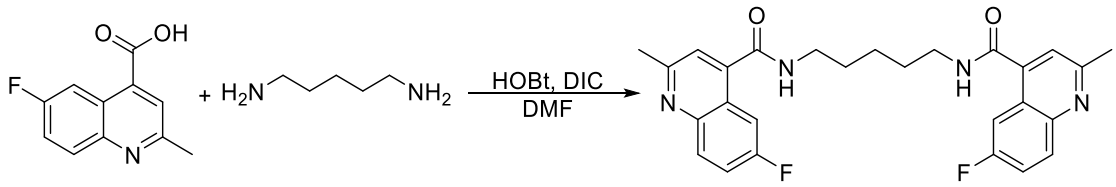
3.18

3.18 Pale yellow powder	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 10.80 (s, 2 H), 8.12-8.09 (m, 2 H), 8.02 (d, *J* = 8.1 Hz, 2 H), 7.81 (s, 4 H), 7.80-7.77 (m, 2 H), 7.65 (s, 2 H), 7.62 (ddd, *J* = 8.3, 7.0, 1.1 Hz, 2 H), 2.74 (s, 6 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 164.9, 158.4, 156.6, 147.5, 138.9, 134.8, 129.6, 126.4, 122.2, 120.5 (2 C), 120.2, 23.1
	IR	(υ_{max}/cm^{-1}) 3213, 3132, 2970, 2843, 1616, 1500, 1462, 1381
	MS	Found 446.9 $[M+H]^+$, Calculated for $C_{28}H_{22}N_4O_2$ 447.51 $[M+H]^+$
	HRMS	Observed 447.1832 $[M+H]^+$; Theoretical 447.1816 $[C_{28}H_{22}N_4O_2+H]^+$
	Yield	69 mg (55% yield)

A1.3 Synthesis of *bis*-(6-fluoro-2-methylquinoline) molecules in library-2B

Synthesis of N,N'-(pentane-1,5-diyl)bis(6-fluoro-2-methylquinoline-4-carboxamide) (3.20)

To a stirred solution of 6-fluoro-2-methylquinoline-4-carboxylic acid (50 mg, 0.24 mmol) and DIC (66 μL, 0.43 mmol) in dimethylformamide (5 mL) was added HOBt (66 mg, 0.49 mmol) in one portion at room temperature. After 30 min, 1,5-diaminopentane (17 mg, 0.12 mmol) was added and the mixture was left to react for 18 h. The reaction was quenched with water (10 mL) and extracted with ethyl acetate (3x20 mL). Then the combined organic phase was washed with a saturated aqueous solution of sodium bicarbonate (30 mL), dried over magnesium sulphate, filtered and concentrated to give the compound **3.20** as pale yellow solid after purification *via* flash column chromatography using methanol/ethyl acetate/dichloromethane 1:2:97 followed by recrystallization with diethyl ether and dichloromethane and the desired product **3.20** was obtained as pale yellow powder.

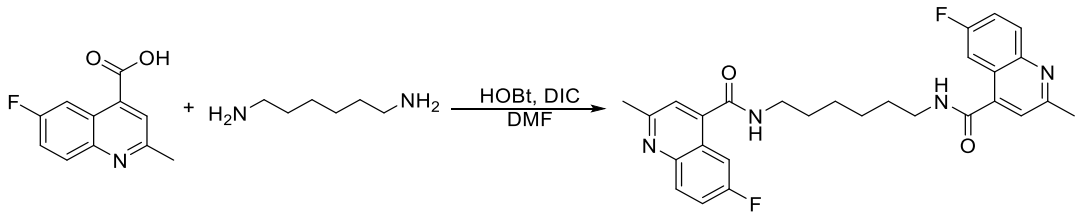


3.20

3.20 Pale yellow powder	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.63 (dd, *J* = 8.6, 5.1 Hz, 2 H), 8.29-8.26 (m, 2 H), 7.94-7.86 (m, 2 H), 7.69-7.66 (m, 2 H), 7.42-7.34 (m, 2 H), 3.16 (t, *J* = 6.8 Hz, 4 H), 2.46 (s, 6 H), 1.60-1.56 (m, 4 H), 1.32-1.28 (m, 2 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 166.4, 158.3, 156.9, 145.1, 142.1, 131.6, 126.6, 123.5, 120.9, 107.2, 40.0, 29.1, 23.5, 23.4
	IR	(υ_{max}/cm^{-1}) 3336, 2931, 2874, 1612, 1558, 1523, 1462, 1385, 1361
	MS	Found 477.1 $[M+H]^+$, Calculated for $C_{27}H_{26}F_2N_4O_2$ 477.2 $[M+H]^+$
	HRMS	Observed 477.2101 $[M+H]^+$; Theoretical 477.2097 $[C_{27}H_{26}F_2N_4O_2+H]^+$
	Yield	32 mg (37% yield)

Synthesis of *N,N′*-(hexane-1,6-diyl)*bis*(6-fluoro-2-methylquinoline-4-carboxamide) (3.21)

Method described for synthesis of **3.20** was employed.

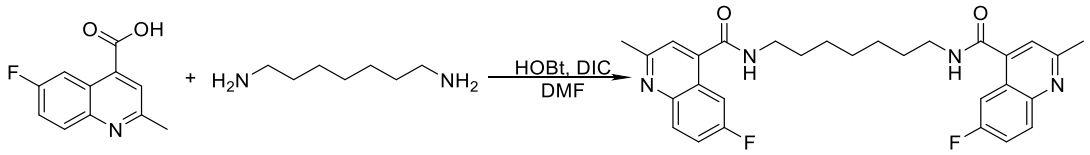


3.21

3.21 White solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.79-8.77 (m, 2 H), 8.03 (dd, *J* = 9.1, 5.5 Hz, 2 H), 7.84-7.76 (m, 2 H), 7.72-7.64 (m, 2 H), 7.51-7.48 (m, 2 H), 3.66-3.60 (m, 4 H), 2.71 (s, 6 H), 1.60-1.58 (m, 4 H), 1.43-1.39 (m, 4 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 166.4, 158.3, 156.9, 145.1, 141.7, 131.6, 123.6, 123.5, 120.9, 108.7, 40.3, 29.1, 26.4, 24.8
	IR	(υ_{max}/cm^{-1}) 3275, 2966, 2943, 2854, 1639, 1593, 1562, 1547, 1288
	MS	Found 491.2 $[M+H]^+$, Calculated for $C_{28}H_{28}F_2N_4O_2$ 491.22 $[M+H]^+$
	HRMS	Observed 491.2252 $[M+H]^+$; Theoretical 491.2253 $[C_{28}H_{28}F_2N_4O_2+H]^+$
	Yield	28 mg (31% yield)

Synthesis of *N,N′*-(heptane-1,7-diyl)*bis*(6-fluoro-2-methylquinoline-4-carboxamide) (3.22)

Protocol used to prepare compound **3.20** was followed.

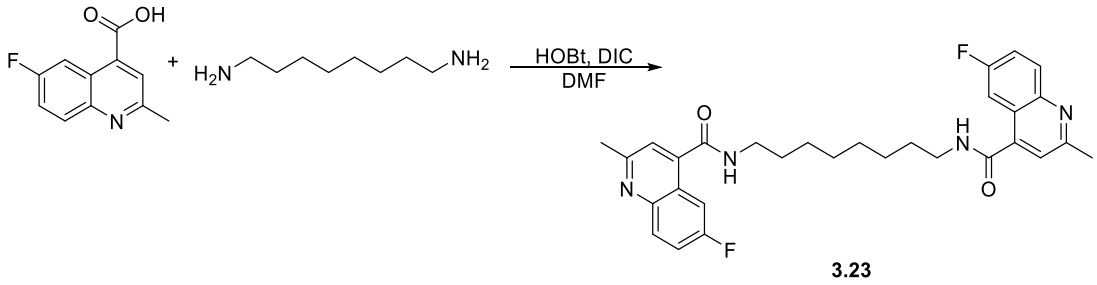


3.22

3.22 White powder	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.93 (dd, *J* = 9.2, 2.7 Hz, 1 H), 8.89 (dd, *J* = 9.2, 2.9 Hz, 1 H), 8.25 (br s, 2 H), 8.08 (dd, *J* = 5.0 Hz, 2.3 Hz, 2 H), 7.71 (s, 2 H), 7.42-7.39 (m, 2 H), 3.39-3.36 (m, 4 H), 2.52 (s, 6 H), 1.53-1.49 (m, 4 H), 1.26-1.23 (m, 6 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 165.3, 162.9, 158.2, 145.3, 142.1, 132.5, 125.4, 122.7, 120.1, 109.8, 41.1, 29.7, 29.1, 26.3, 23.2
	IR	(υ_{max}/cm^{-1}) 3282, 2966, 2965, 2858, 1639, 1593, 1549, 1465
	MS	Found 505.20 $[M+H]^+$, Calculated for $C_{29}H_{30}F_2N_4O_2$ 505.23 $[M+H]^+$
	HRMS	Observed 505.2410 $[M+H]^+$; Theoretical 505.2410 $[C_{29}H_{30}F_2N_4O_2+H]^+$
	Yield	23 mg (25% yield)

Synthesis of *N,N′*-(octane-1,8-diyl)*bis*(6-fluoro-2-methylquinoline-4-carboxamide) (3.23)

Compound **3.23** was obtained as white solid by following the general protocol used for **3.20** after purification *via* flash column chromatography using n-hexane/ethyl acetate (0-70%).

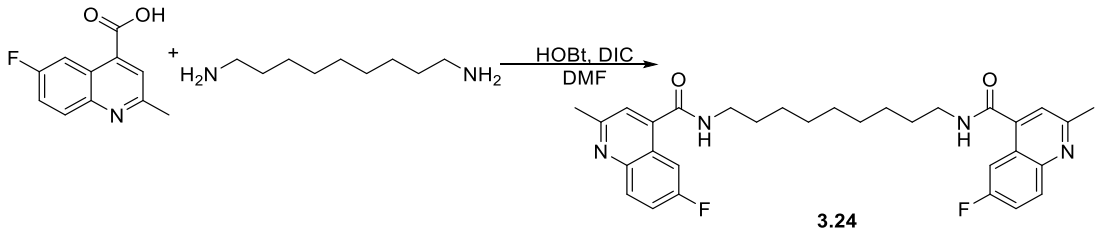


3.23

3.23 White solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.78 (d, *J* = 5.0 Hz, 1 H), 8.03 (dd, *J* = 9.2, 5.7 Hz, 1 H), 7.79 (dd, *J* = 9.2, 5.7 Hz, 1 H), 7.71-7.64 (m, 2 H), 7.52-7.49 (m, 1 H), 5.48 (d, *J* = 6.6 Hz, 2 H), 3.67-3.58 (m, 2 H), 3.34-3.31 (m, 6 H), 2.69-2.66 (m, 2 H), 1.36-1.34 (m, 2 H), 1.02-0.97 (m, 12 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 166.5, 158.5, 157.2, 145.3, 141.9, 131.8, 123.4, 121.1, 120.1, 101.5, 39.3, 29.3, 29.1, 26.8, 23.7
	IR	(υ_{max}/cm^{-1}) 3275, 2966, 2851, 1643, 1593, 1547, 1466, 1334
	MS	Found 519.10 $[M+H]^+$, Calculated for $C_{30}H_{32}F_2N_4O_2$ 519.25 $[M+H]^+$
	HRMS	Observed 519.2567 $[M+H]^+$; Theoretical 519.2566 $[C_{30}H_{32}F_2N_4O_2+H]^+$
	Yield	58 mg (92% yield)

Synthesis of *N,N'*-(nonane-1,9-diyl)*bis*(6-fluoro-2-methylquinoline-4-carboxamide) (3.24)

Compound **3.24** was obtained by following same protocol employed for compound **3.23**.

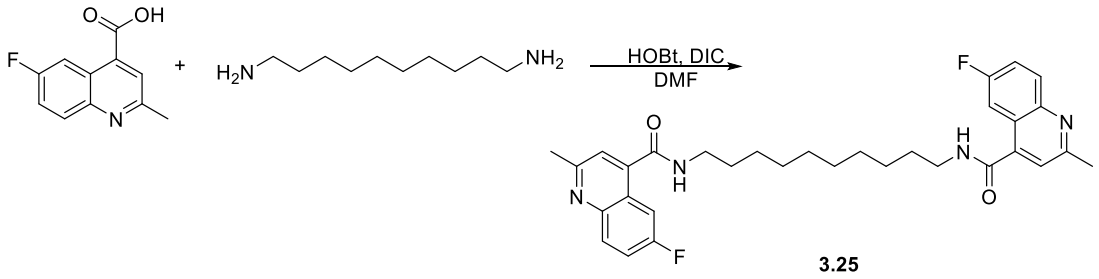


3.24

3.24 White powder	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.77 (t, *J* = 5.4 Hz, 2 H), 8.04 (dd, *J* = 9.2, 5.7 Hz, 2 H), 7.78 (d, *J* = 2.8 Hz, 1 H), 7.80 (d, *J* = 2.8 Hz, 1 H), 7.67 (td, *J* = 8.7, 2.9 Hz, 2 H), 7.50 (s, 2 H) 3.32-3.28 (m, 4 H), 2.67 (s, 6 H), 1.60-1.54 (m, 4 H), 1.34-1.30 (m, 10 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 166.1, 158.3, 158.1, 144.9, 141.6, 131.5, 123.4, 123.2, 120.7, 108.7, 38.9, 29.0, 28.9, 28.7, 26.4, 24.6
	IR	(υ_{max}/cm^{-1}) 3298, 2935, 2854, 1639, 1593, 1539, 1466, 1334
	MS	Found 533.21 $[M+H]^+$, Calculated for $C_{31}H_{34}F_2N_4O_2$ 533.26 $[M+H]^+$
	HRMS	Observed 533.2716 $[M+H]^+$; Theoretical 533.2723 $[C_{31}H_{34}F_2N_4O_2+H]^+$
	Yield	52 mg (81% yield)

Synthesis of *N,N'*-(decane-1,10-diyl)*bis*(6-fluoro-2-methylquinoline-4-carboxamide) (3.25)

Protocol used to prepare compound **3.23** was followed.



3.25

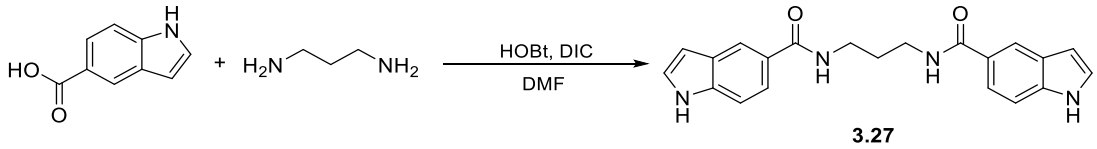
3.25 White powder	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 8.03 (dd, *J* = 9.3, 5.3 Hz, 2 H), 7.79 (dd, *J* = 9.8, 2.8 Hz, 3 H), 7.67 (d, *J* = 8.3 Hz, 1 H), 7.47 (ddd, *J* = 9.3, 7.9, 2.9 Hz, 2 H), 6.33 (t, *J* = 5.7 Hz, 2 H), 7.31 (s, 2 H), 3.56-3.51 (m, 3 H), 2.70 (s, 6 H), 1.70-1.65 (m, 4 H), 1.46-1.39 (m, 5 H), 1.36-1.32 (m, 4 H), 1.26 (s, 2 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 175.1, 168.3, 154.1, 144.8, 143.5, 130.0, 128.9, 120.8, 120.1, 104.4, 40.1, 32.0, 29.4, 28.9, 26.7, 24.8
	IR	(υ_{max}/cm^{-1}) 3271, 2920, 2850, 1643, 1547, 1226, 1161
	MS	Found 547.30 $[M+H]^+$, Calculated for $C_{32}H_{36}F_2N_4O_2$ 547.28 $[M+H]^+$

	HRMS	Observed 547.2885 $[M+H]^+$; Theoretical 547.2879 $[C_{32}H_{36}F_2N_4O_2+H]^+$
	Yield	60 mg (90% yield)

A1.4 Synthesis of *bis*-indole compounds in library-2C

Synthesis of *N,N*'-(propane-1,3-diyl)bis(1*H*-indole-5-carboxamide) (3.27)

To a suspension of 1*H*-indole-5-carboxylic acid (25 mg, 0.16 mmol) in dimethylformamide (5 mL), HOBt (42 mg, 0.31 mmol) and DIC (42 μl, 0.27 mmol) were added in one portion at room temperature. After 30 minutes 1,3-diaminopropane (6 mg, 6 μL, 0.08 mmol) was added and the reaction mixture was left under stirring for 24 h, quenched with water (5 mL). Phases were separated, the organic layer was washed with a saturated aqueous solution of sodium bicarbonate, dried over magnesium sulphate, filtered and concentrated. Purification was carried out by flash column chromatography with n-hexane/ethyl acetate (0-70%).

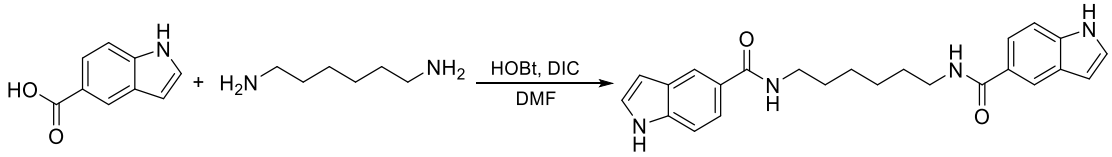


3.27

3.27 Pale pink solid	**^{1}H NMR**	(400 MHz, MeOD) δ 8.33 (dd, *J* = 1.6, 0.6 Hz, 2 H), 7.80 (dd, *J* = 8.6, 1.8 Hz, 2 H), 7.43-7.39 (m, 2 H), 7.32 (d, *J* = 3.3 Hz, 2 H), 6.57 (dd, *J* = 3.2, 0.9 Hz, 2 H), 3.68-3.64 (m, 4 H), 1.10-1.02 (m, 2 H)
	^{13}C NMR	(100 MHz, MeOD) δ 166.2, 140.7, 129.4, 127.6, 125.0, 124.2, 122.7, 112.2, 104.2, 48.7, 33.2
	IR	(υ_{max}/cm^{-1}) 3352, 2655, 1605, 1578, 1489, 1462, 1435, 1385
	MS	Found 361.27 $[M+H]^+$, Calculated for $C_{21}H_{20}N_4O_2$ 360.42 $[M+H]^+$
	HRMS	Observed 362.1684 $[M+H]^+$, Theoretical 362.1689 $[C_{21}H_{20}N_4O_2+H]^+$
	Yield	26 mg (93% yield)

Synthesis of *N,N'*-(hexane-1,6-diyl)bis(1*H*-indole-5-carboxamide) (3.28)

Method described for preparation of **3.27** was employed. Desired compound **3.28** was obtained as a tan solid after purification by flash column chromatography using ethyl acetate/methanol (0-2%).

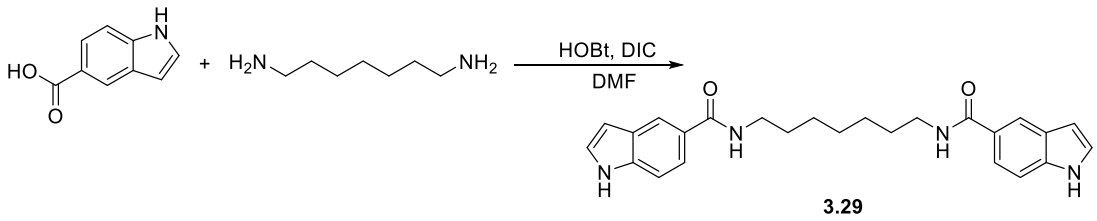


3.28

3.28 Tan powder	**^{1}H NMR**	(400 MHz, MeOD) δ 8.31 (dd, *J* = 1.7, 0.6 Hz, 2 H), 7.81 (dd, *J* = 8.4, 1.7 Hz, 2 H), 7.44-7.40 (m, 2 H), 7.15 (d, *J* = 1.5 Hz, 2H), 6.59 (dd, *J* = 3.2, 0.9 Hz, 2 H), 3.48 (t, *J* = 6.5 Hz, 4 H), 1.66-1.62 (m, 4 H), 1.29-1.27 (m, 2 H)
	^{13}C NMR	(100 MHz, MeOD) δ 168.2, 139.8, 129.5, 126.6, 121.9, 119.2, 112.4, 110.6, 103.1, 41.2, 29.2, 24.8
	IR	(υ_{max}/cm^{-1}) 3450, 3201, 2952, 2864, 1635, 1601, 1578, 1443, 1431
	MS	Found 403.10 $[M+H]^+$, Calculated for $C_{24}H_{26}N_4O_2$ 403.50 $[M+H]^+$
	HRMS	Observed 403.2119 $[M+H]^+$; Theoretical 403.2129 $[C_{24}H_{26}N_4O_2+H]^+$
	Yield	26 mg (92% yield)

Synthesis of *N,N*'-(heptane-1,7-diyl)bis(1*H*-indole-5-carboxamide) (3.29)

Method described for preparation of **3.27** was employed. Desired compound **3.29** obtained as a tan solid after purification done by flash column chromatography using ethyl acetate/methanol (0-3%).

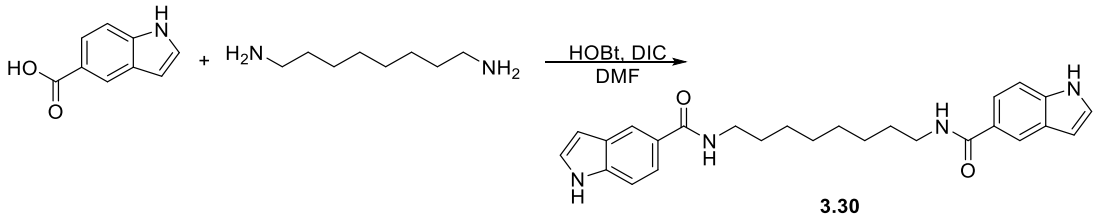


3.29

3.29 Tan solid	**^{1}H NMR**	(400 MHz, MeOD) δ 8.66 (dd, *J* = 1.6, 0.6 Hz, 2 H), 7.83 (dd, *J* = 8.5, 1.7 Hz, 2 H), 7.37 (d, *J* = 3.3 Hz, 2 H), 7.17 (d, *J* = 1.4 Hz, 2 H), 6.61 (dd, *J* = 3.1, 0.9 Hz, 2 H), 3.41 (t, *J* = 6.1 Hz, 4 H), 1.62-1.54 (m, 6 H), 1.26-1.21 (m, 4 H)
	^{13}C NMR	(100 MHz, MeOD) δ 166.7, 141.5, 130.3, 127.4, 119.6, 118.7, 111.8, 109.7, 103.8, 40.4, 30.7, 29.2, 26.4
	IR	(υ_{max}/cm^{-1}) 3202, 2922, 2868, 1601, 1531, 1300
	MS	Found 417.10 $[M+H]^+$, Calculated for $C_{25}H_{28}N_4O_2$ 417.52 $[M+H]^+$
	HRMS	Observed 417.2276 $[M+H]^+$; Theoretical 417.2285 $[C_{25}H_{28}N_4O_2+H]^+$
	Yield	21 mg (73% yield)

Synthesis of *N,N'*-(octane-1,8-diyl)bis(1*H*-indole-5-carboxamide) (3.30)

Method described for preparation of **3.27** was employed. Desired compound **3.30** obtained as amber oil after purification by flash column chromatography using ethyl acetate/methanol (0-2.5%).

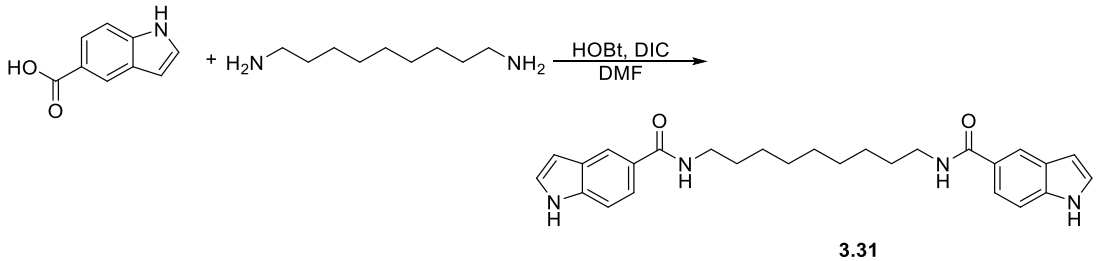


3.30

3.30 Amber oil	**^{1}H NMR**	(400 MHz, MeOD) δ 8.96 (dd, *J* = 1.8, 0.8 Hz, 2 H), 8.01 (dd, *J* = 8.5, 1.7 Hz, 2 H), 7.82 (d, *J* = 3.3 Hz, 2 H), 7.14 (d, *J* = 1.5 Hz, 2 H), 6.58 (d, *J* = 3.1 Hz, 2 H), 3.36 (t, *J* = 6.2 Hz, 4 H), 1.56 (t, *J* = 7.6 Hz, 4 H), 1.28-1.26 (m, 8 H)
	^{13}C NMR	(100 MHz, MeOD) δ 164.1, 140.3, 129.4, 126.5, 120.2, 118.7, 112.3, 110.0, 103.4, 39.9, 30.2, 29.8, 26.7
	IR	(υ_{max}/cm^{-1}) 3250, 2976, 2849, 1605, 1574, 1469, 1392, 1373
	MS	Found 431.20 $[M+H]^+$, Calculated for $C_{26}H_{30}N_4O_2$ 431.55 $[M+H]^+$
	HRMS	Observed 431.2430 $[M+H]^+$; Theoretical 431.2442 $[C_{26}H_{30}N_4O_2+H]^+$
	Yield	25 mg (94% yield)

Synthesis of *N,N'*-(nonane-1,9-diyl)bis(1*H*-indole-5-carboxamide) (3.31)

Method described for preparation of **3.27** was employed. Desired compound **3.31** obtained as an amber oil after purification by flash column chromatography using ethyl acetate/methanol (0-70%).



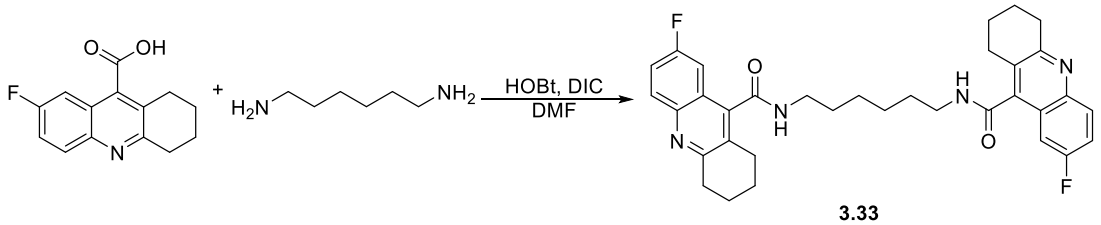
3.31

3.31 Amber oil	**^{1}H NMR**	(400 MHz, MeOD) δ 8.78 (dd, *J* = 1.8, 0.8 Hz, 2 H), 8.34 (dd, *J* = 8.5, 1.7 Hz, 2 H), 7.95 (d, *J* = 3.3 Hz, 2 H), 7.86 (d, *J* = 1.5 Hz, 2 H), 6.58 (d, *J* = 3.1 Hz, 2 H) 3.32 (t, *J* = 6.2 Hz, 4 H) 1.59-1.55 (m, 4 H) 1.29-1.25 (m, 10 H)
	^{13}C NMR	(100 MHz, MeOD) δ 159.9, 139.3, 131.2, 127.0, 121.3, 119.6, 112.6, 110.2, 102.0, 41.0, 30.5, 29.4 (2 C), 25.9
	IR	(υ_{max}/cm^{-1}) 3252, 2928, 2854, 1624, 1574, 1466, 1388
	MS	Found 445.20 $[M+H]^+$, Calculated for $C_{27}H_{32}N_4O_2$ 445.58 $[M+H]^+$
	HRMS	Observed 445.2587 $[M+H]^+$; Theoretical 445.2598 $[C_{27}H_{32}N_4O_2+H]^+$
	Yield	24 mg (89% yield)

A1.5 Synthesis of *bis*-acridine molecules in library-3A

Synthesis of *N,N*'-(hexane-1,6-diyl)*bis*(7-fluoro-1,2,3,4-tetrahydroacridine-9-carboxamide) (3.33)

To a stirred solution of 7-fluoro-1,2,3,4-tetrahydroacridine-9-carboxylic acid (75 mg, 0.31 mmol) and DIC (89 μL, 0.57 mmol) in dimethylformamide (8 mL), HOBt (83 mg, 0.61 mmol) was added in one portion at room temperature and stirred for 30 minutes. Then 1,8-diaminohexane (22 mg, 0.15 mmol) was added and allowed to react for 14 h. The reaction was quenched with water and extracted with ethyl acetate (3x30 mL), organics were combined and washed with a saturated aqueous solution of sodium bicarbonate (30 mL), dried over magnesium sulphate, filtered and concentrated. Purification by flash column chromatography with ethyl acetate/methanol (0-2.5%).

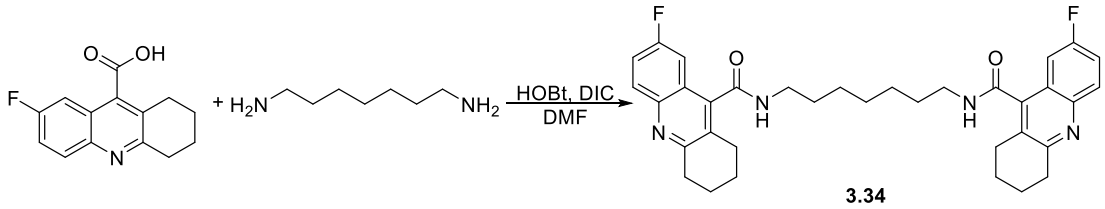


3.33

3.33 Tan solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.72 (t, *J* = 5.5 Hz, 2 H), 7.98 (dd, *J* = 9.3, 5.5 Hz, 2 H), 7.60 (td, *J* = 8.8, 2.8 Hz, 2 H), 7.28 (dd, *J* = 9.9, 2.9 Hz, 2 H), 3.40-3.35 (m, 4 H), 3.05-3.00 (m, 4 H), 2.85 (t, *J* = 6.2 Hz, 4 H), 1.89 (d, *J* = 5.5 Hz, 4 H), 1.85-1.80 (m, 4 H), 1.59 (d, *J* = 6.0 Hz, 4 H), 1.44 (br s, 4 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 165.8, 158.4 (2 C), 142.9, 141.8, 131.2, 126.9, 121.1, 117.1, 107.5, 40.6, 33.2, 28.9, 28.2, 22.3, 22.0 (2 C)
	IR	(υ_{max}/cm^{-1}) 3232, 3055, 2947, 2866, 1632, 1555, 1496, 1450
	MS	Found 571.2 $[M+H]^+$, Calculated for $C_{34}H_{36}F_2N_4O_2$ $[M+H]^+$ 571.28
	HRMS	Observed 571.2877 $[M+H]^+$; Theoretical 571.2879 $[C_{34}H_{36}F_2N_4O_2+H]^+$
	Yield	47 mg (55% yield)

Synthesis of *N,N'*-(heptane-1,7-diyl)*bis*(7-fluoro-1,2,3,4-tetrahydroacridine-9-carboxamide) (3.34)

Protocol described in the synthesis of **3.33** was used. Purification was carried out by flash column chromatography with ethyl acetate/methanol (0-4%).

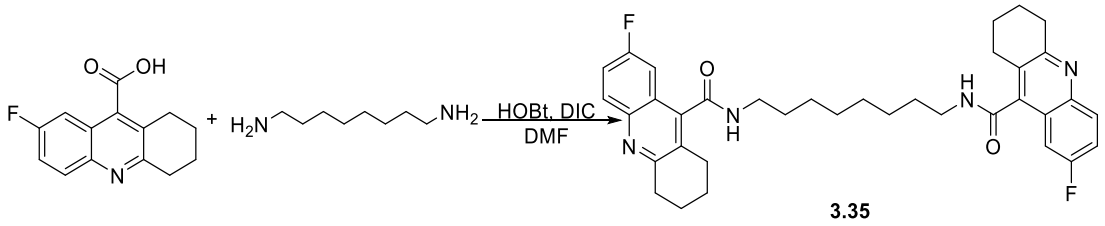


3.34

3.34 Pale yellow solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.73-8.69 (m, 2 H), 7.98 (dd, *J* = 9.2, 5.7 Hz, 2 H), 7.60 (td, *J* = 8.7, 2.5 Hz, 2 H), 7.28 (dd, *J* = 9.8, 2.5 Hz, 2 H), 3.05-3.00 (m, 4 H), 2.86-2.82 (m, 4 H), 1.87-1.84 (m, 4 H), 1.85-1.78 (m, 6 H), 1.61-1.56 (m, 4 H), 1.40-1.38 (m, 8 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 165.8, 164.4 (2 C), 142.9, 141.8, 131.3, 126.9, 123.7, 119.0, 107.7, 40.1, 33.2, 28.9, 28.4, 26.5, 26.4, 22.2 (2 C)
	IR	(υ_{max}/cm^{-1}) 3225, 3074, 2928, 2858, 1631, 1558, 1493, 1454, 1431
	MS	Found 585.30 $[M+H]^+$, Calculated for $C_{35}H_{38}F_2N_4O_2$ 585.30 $[M+H]^+$
	HRMS	Observed 585.3028 $[M+H]^+$; Theoretical 585.3036 $[C_{35}H_{38}F_2N_4O_2+H]^+$
	Yield	34 mg (39% yield)

Synthesis of *N,N'*-(octane-1,8-diyl)*bis*(7-fluoro-1,2,3,4-tetrahydroacridine-9-carboxamide) (3.35)

Protocol used in the synthesis of compound **3.33** was employed and desired product was obtained after purification done by flash column chromatography with ethyl acetate/methanol (0-2%).



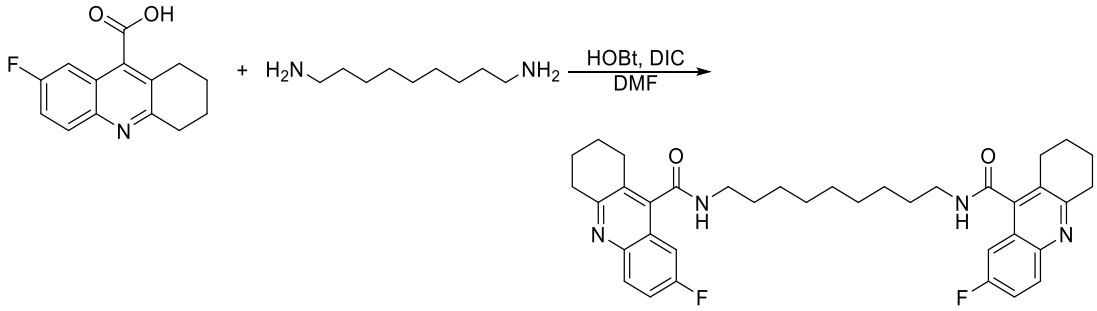
3.35

3.35 Yellow powder	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 8.70 (t, *J* = 5.5 Hz, 2 H), 7.98 (dd, *J* = 9.3, 5.5 Hz, 2 H), 7.60 (td, *J* = 8.8, 2.8 Hz, 2 H), 7.27 (dd, *J* = 9.8, 2.8 Hz, 2 H), 3.66-3.59 (m, 2 H), 3.40-3.36 (m, 2 H), 3.05-3.00 (m, 4 H), 2.84 (t, *J* = 6.4 Hz, 4 H), 1.89 (d, *J* = 5.0 Hz, 4 H), 1.84-1.79 (m, 4 H), 1.57 (d, *J* = 6.8 Hz, 4 H), 1.42-1.37 (m, 8 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 168.2, 164.4 (2 C), 142.9, 141.3, 131.9, 126.6, 123.3, 120.0, 105.1, 39.6, 33.8, 31.4, 27.5, 27.1, 26.4, 21.9 (2 C)
	IR	IR (υ_{max}/cm^{-1}) 3336, 3271, 2858, 1628, 1554, 1493, 1454, 1442
	MS	Found 599.30 $[M+H]^+$, Calculated for $C_{36}H_{40}F_2N_4O_2$ 599.31 $[M+H]^+$

	HRMS	Observed 599.3190 $[M+H]^+$; Theoretical 599.3192 $[C_{36}H_{40}F_2N_4O_2+H]^+$
	Yield	52 mg (57% yield)

Synthesis of *N,N'*-(nonane-1,9-diyl)*bis*(7-fluoro-1,2,3,4-tetrahydroacridine-9-carboxamide) (3.36)

Compound **3.36** was prepared by following protocol describer in the synthesis of **3.33** and purified by flash column chromatography with ethyl acetate/methanol (0-10%).



3.36

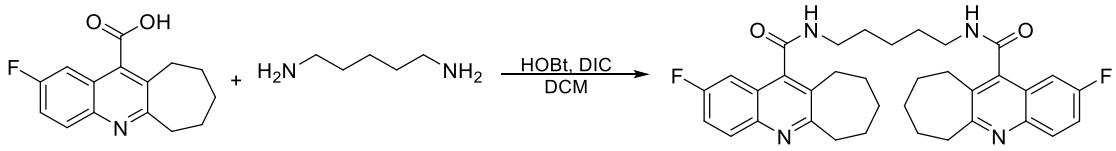
3.36 White solid	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 7.89 (dd, *J* = 9.2, 5.4 Hz, 2 H), 7.35 (td, *J* = 8.6, 2.6 Hz, 2 H), 7.27-7.22 (m, 2 H), 6.40 (t, *J* = 5.7 Hz, 2 H), 3.52 (q, *J* = 6.6 Hz, 4 H), 3.02-2.98 (m, 4 H), 2.83 (t, *J* = 5.2 Hz, 4 H), 1.91-1.85 (m, 4 H), 1.83-1.77 (m, 4 H), 1.69-1.63 (m, 4 H), 1.45-1.36 (m, 10 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 167.6, 165.0 (2 C), 144.1, 139.2, 133.6, 127.3, 122.9, 121.1, 103.4, 40.6, 32.8, 29.8, 27.1, 26.4, 26.2, 22.2 (2 C)
	IR	(υ_{max}/cm^{-1}) 3336, 2935, 2874, 1612, 1562, 1523, 1462, 1384, 1327
	MS	Found 613.30 $[M+H]^+$, Calculated for $C_{37}H_{42}F_2N_4O_2$ 613.33 $[M+H]^+$
	HRMS	Observed 613.3354 $[M+H]^+$; Theoretical 613.3349 $[C_{37}H_{42}F_2N_4O_2+H]^+$
	Yield	70 mg (74% yield)

A1.6 Synthesis of *bis*- (2-fluoro-7,8,9,10-tetrahydro-6H-cyclohepta[b] quinoline-11-carboxamide in library-3B

Synthesis of *N,N'*-(pentane-1,5-diyl)bis(2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*] quinoline-11-carboxamide) (3.38)

A stirred suspension of 2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxylic acid (124 mg, 0.48 mmol) in dichloromethane (8 mL) was treated with HOBt (129 mg, 0.96 mmol) and DIC (130 μl, 0.84 mmol) at room temperature. After 30 min, 1,5-diaminopentane (25 mg, 0.20 mmol) was added, and left to react for 6 h. Reaction progress was monitored by TLC and additional 49 mg (0.19 mmol) of acid, 51 μL (0.33 mmol) of DIC and 51 mg (0.38 mmol) of HOBt were added and the resulting mixture was stirred for 16 h, then quenched with water (10 mL) and extracted with

ethyl acetate (2x30 mL). Organics were combined, washed with a saturated aqueous solution of sodium bicarbonate (20 mL), dried over magnesium sulphate, filtered and concentrated. As the concentrated product has got some impurities, finally purification was carried out by 'Catch and release method' using SCX-2 cartridge (10 g) to get the compound **3.38**.

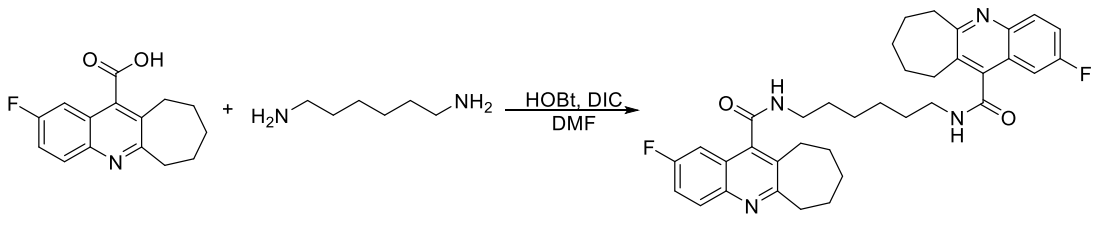


3.38

3.38 Tan solid	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 8.01-7.96 (m, 2 H), 7.93 (dd, *J* = 9.3, 5.3 Hz, 2 H), 7.36 (dd, *J* = 9.3, 2.6 Hz, 2 H), 6.05-5.99 (m, 2 H), 3.60 (d, *J* = 7.3 Hz, 4 H), 3.24-3.17 (m, 4 H), 2.96-2.88 (m, 4 H), 2.77-2.73 (m, 2 H), 1.92-1.86 (m, 4 H), 1.80 (d, *J* = 6.3 Hz, 4 H), 1.25 (d, *J* = 6.8 Hz, 4 H), 1.18 (d, *J* = 6.5 Hz, 2 H), 1.15 (d, *J* = 6.3 Hz, 2 H).
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 168.7, 165.7, 156.7, 142.9, 137.8, 131.9, 128.2, 126.8, 123.9, 101.2, 46.9, 37.2, 32.5, 30.4 (2 C), 26.6 (2 C), 24.0
	IR	(υ_{max}/cm^{-1}) 3258, 2925, 2854, 1711, 1626, 1548, 1499, 1385
	MS	Found 583.00 $[M-H]^+$, Calculated for $C_{35}H_{38}F_2N_4O_2$ 584.70 $[M+H]^+$
	HRMS	Observed 584.3114 $[M+H]^+$, Theoretical 584.3102 $[C_{35}H_{38}F_2N_4O_2+H]^+$
	Yield	72 mg (84% yield)

Synthesis of *N,N'*-(hexane-1,6-diyl)bis(2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxamide) (3.39)

Compound **3.39** was prepared with the method described for the synthesis of compound **3.38** using dimethylformamide as solvent.



3.39

3.39 Amber solid	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 8.68 (dd, *J* = 8.3, 0.8 Hz, 2 H), 8.07-8.03 (m, 2 H), 7.97-7.93 (m, 2 H), 7.70 (td, *J* = 7.8, 0.8 Hz, 2 H), 3.78 (spt, *J* = 6.5 Hz, 4 H), 3.30-3.25 (m, 4 H), 2.84 (dd, *J* = 6.3, 4.5 Hz, 4 H), 1.89-1.86 (m, 4 H), 1.85-1.81 (m, 4 H), 1.81-1.75 (m, 4 H), 1.71-1.65 (m, 4 H), 1.52 (br s, 4 H)

	^{13}C NMR	(100 MHz, $CDCl_3$) δ 164.6, 162.4, 159.9, 142.9, 136.1, 131.9, 128.2, 126.8, 123.9, 107.6, 42.6, 32.7, 29.9 (2 C), 27.7, 26.8 (2 C), 23.6
	IR	(υ_{max}/cm^{-1}) 3337, 2966, 2925, 2850, 1727, 1615, 1566, 1458, 1380
	MS	Found 598.82 $[M+H]^+$, Calculated for $C_{36}H_{40}F_2N_4O_2$ 598.73 $[M+H]^+$
	HRMS	Observed 599.3185 $[M+H]^+$, Theoretical 599.3192 $[C_{36}H_{40}F_2N_4O_2+H]^+$
	Yield	24 mg (73% yield)

Synthesis of *N,N'*-(heptane-1,7-diyl)bis(2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxamide) (3.40)

Protocol described in the synthesis of **3.38** was used. Purification was carried out by flash column chromatography with ethyl acetate/methanol (0-1%) to give the product **3.40** as pale yellow solid.

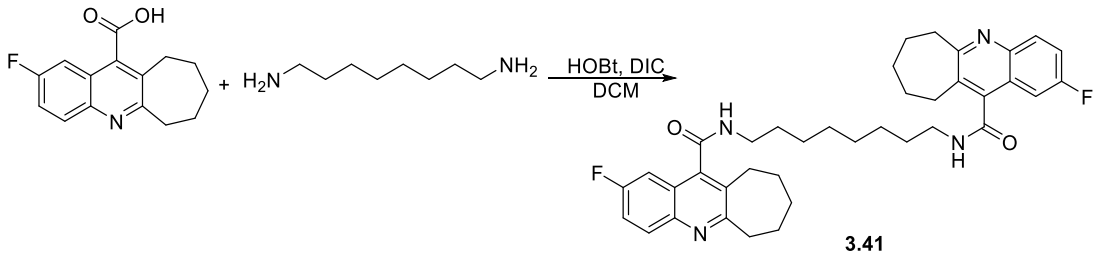
3.40

3.40 Pale orange solid	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 7.97 (dd, *J* = 9.1, 5.6 Hz, 2 H), 7.40-7.36 (m, 2 H), 7.34-7.32 (m, 2 H), 5.91 (t, *J* = 6.4 Hz, 2 H), 3.88-3.83 (m, 4 H), 3.63-3.57 (m, 4 H), 2.97-2.85 (m, 4 H), 1.91-1.85 (m, 4 H), 1.78-1.70 (m, 8 H), 1.51-1.49 (m, 4 H), 1.16-1.12 (m, 6 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 167.7, 164.9, 163.2, 148.7, 147.9, 132.1, 127.3, 122.6, 119.4, 101.4, 42.6, 36.6, 35.3, 29.2, 27.9, 26.4, 25.2, 24.3, 23.8
	IR	(υ_{max}/cm^{-1}) 3253, 2927, 2847, 1634, 1550, 1497, 1455, 1385
	MS	Found 611.01 $[M-H]^+$, Calculated for $C_{37}H_{42}F_2N_4O_2$ 612.33 $[M+H]^+$
	HRMS	Observed 613.3268 $[M+H]^+$, Theoretical 613.3286 $[C_{37}H_{42}F_2N_4O_2+H]^+$
	Yield	111 mg (94% yield)

Synthesis of *N,N'*-(octane-1,8-diyl)bis(2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxamide) (3.41)

Method described for the synthesis of **3.38** was employed. The crude was purified by flash column chromatography using n-hexane/ethyl acetate 0-40%) to afford the desired compound as white solid.

.

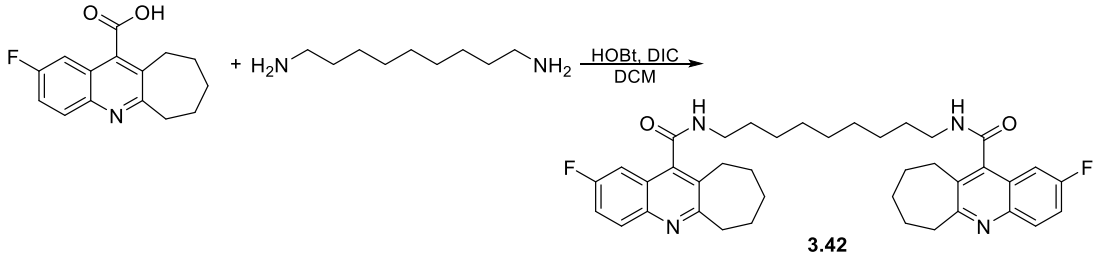


3.41

3.41 White solid	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 7.99 (dd, *J* = 9.1, 5.3 Hz, 2 H), 7.38-7.34 (m, 2 H), 7.34-7.31 (m, 2 H), 6.14 (t, *J* = 5.6 Hz, 2 H), 3.57 (q, *J* = 6.7 Hz, 4 H), 3.20-3.15 (m, 4 H), 2.92-2.87 (m, 4 H), 1.87 (dt, *J* = 11.1, 5.8 Hz, 4 H), 1.75-1.70 (m, 8 H), 1.44 (s, 4 H), 1.13 (d, *J* = 6.3 Hz, 8 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 167.4, 164.2, 158.9, 149.4, 148.0, 130.2, 127.3, 122.6, 120.1, 107.4, 42.3, 39.8, 37.7, 31.8, 29.5, 28.9, 28.2, 27.4, 26.7
	IR	(υ_{max}/cm^{-1}) 2927, 2853, 2364, 1624, 1559, 1499, 1458
	MS	Found 627.05 $[M+H]^+$, Calculated for $C_{38}H_{44}F_2N_4O_2$ 627.78 $[M+H]^+$
	HRMS	Observed 627.3490 $[M+H]^+$, Theoretical 627.3505 $[C_{38}H_{44}F_2N_4O_2+H]^+$
	Yield	19 mg (28% yield)

Synthesis of *N,N'*-(nonane-1,9-diyl)bis(2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxamide) (3.42)

Compound **3.42** was obtained by following protocol employed to synthesised **3.38**.



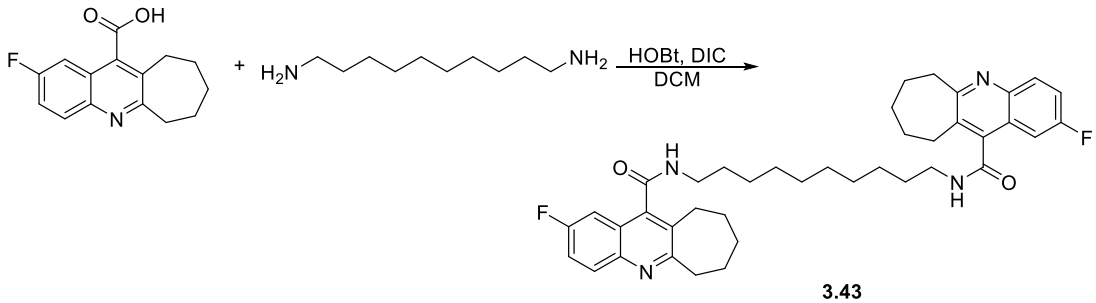
3.42

3.42 Tan solid	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 8.03-7.99 (m, 2 H), 7.65 (d, *J* = 8.3 Hz, 1 H), 7.42-7.38 (m, 1 H), 7.36-7.31 (m, 2 H), 6.52 (t, *J* = 5.8 Hz, 2 H), 3.79 (q, *J* = 6.4 Hz, 4 H), 3.57 (q, *J* = 6.4 Hz, 4 H), 3.19-3.14 (m, 2 H), 2.96 (s, 2 H), 2.88 (s, 2 H), 1.71-1.66 (m, 4 H), 1.13 (m, 20 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 169.2, 164.3, 161.9, 157.2, 147.7, 141.9, 136.4, 126.5, 125.3, 110.4, 41.2, 31.7, 29.3 (2 C), 28.8, 26.7, 26.1 (2 C), 24.1, 23.4
	IR	(υ_{max}/cm^{-1}) 3338, 2928, 2966, 2358, 1614, 1497, 1461, 1384
	MS	Found 641.17 $[M+H]^+$, Calculated for $C_{39}H_{46}F_2N_4O_2$ 641.80 $[M+H]^+$

	HRMS	Observed 641.3639 $[M+H]^+$, Theoretical 641.3662 $[C_{39}H_{46}F_2N_4O_2+H]^+$
	Yield	85 mg (70% yield)

Synthesis of *N,N'*-(decane-1,10-diyl)bis(2-fluoro-7,8,9,10-tetrahydro-6*H*-cyclohepta[*b*]quinoline-11-carboxamide) (3.43)

Protocol described in the synthesis of **3.38** was used and desired compound was obtained after purification by flash column chromatography using n-hexane/ethyl acetate (0-100%).



3.43

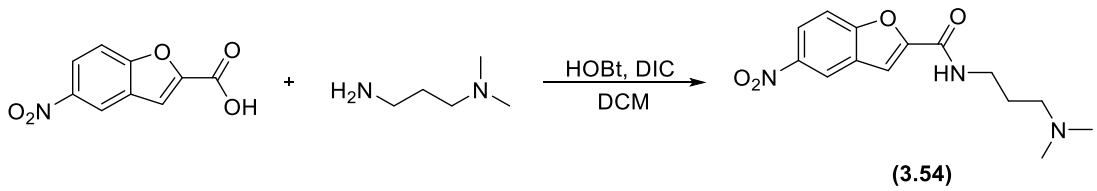
3.43 Pale yellow solid	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 7.99 (dd, *J* = 9.3, 5.5 Hz, 2 H), 7.41 (dd, *J* = 8.6, 3.3 Hz, 2 H), 7.34 (dd, *J* = 9.3, 2.6 Hz, 2 H), 5.91 (t, *J* = 5.5 Hz, 2 H) 3.57 (q, *J* = 6.7 Hz, 4 H), 3.20 (d, *J* = 11.4 Hz, 4 H), 2.94-2.89 (m, 4 H), 1.92-1.87 (m, 4 H), 1.71-1.66 (m, 4 H), 1.41-1.37 (m, 16 H), 1.16 (m, 4 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 168.7, 156.8, 155.1, 147.5, 143.5, 133.3, 127.5, 122.5, 121.8, 115.3, 39.4, 37.1, 34.5, 30.3 (2 C), 29.7, 28.1, 26.8 (2 C), 24.6
	IR	(υ_{max}/cm^{-1}) 3254, 2926, 2852, 1636, 1539, 1498, 1458, 1385
	MS	Found 655.02 $[M+H]^+$, Calculated for $C_{40}H_{48}F_2N_4O_2$ 655.83 $[M+H]^+$
	HRMS	Observed 655.3795 $[M+H]^+$, Theoretical 655.3818 $[C_{40}H_{48}F_2N_4O_2+H]^+$
	Yield	31 mg (27% yield)

A1.7 Synthesis of *bis*-benzofuran type molecules in library-4

[2,2'-Bipyridine]-5,5'-dicarboxylic acid (**3.44**), [1,1'-biphenyl]-4,4'-dicarboxylic acid (**3.45**), [2,2'-bipyridine]-4,4'-dicarboxylic acid (**3.46**) and 5-nitrobenzofuran-2-carboxylic acid (**3.47**) were used to synthesize this library molecule. Additionally six tertiary amine tails, 3-(dimethylamino)-1-propylamine (**3.48**), 1-(3-aminopropyl)pyrrolidine (**3.49**), *N*-(3-aminopropyl)piperidine (**3.50**), *N,N*-dimethyl-1,4-butanediamine (**3.51**), 1-(4-aminobutyl)pyrrolidine (**3.52**) and 4-piperidin-1-yl-butylamine (**3.53**) were used as terminal tail. Final products were prepared in three steps: starting with the synthesis of monoamide intermediate, followed by reduction of the nitro group to generate the corresponding amine which was used to generate the final *bis*-amide product.

Synthesis of *N*-(3-(dimethylamino)propyl)-5-nitrobenzofuran-2-carboxamide (3.54)

To a suspension of 5-nitrobenzofuran-2-carboxylic acid (200 mg, 0.96 mmol) in dichloromethane (10 mL), HOBt (186 mg, 1.40 mmol) and DIC (189 μl, 1.20 mmol) were added. After 30 minutes, 3-(dimethylamino)-1-propylamine (86 μL, 0.70 mmol) was added and the reaction mixture was left under stirring for 5 h at room temperature. The reaction mixture was passed through a SCX-2 cartridge (2 g) to afford the desired compound as an amber oil.

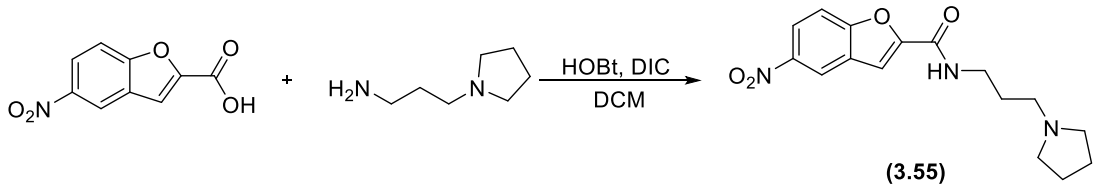


(3.54)

3.54 Amber oil	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 8.78 (s, 1 H), 8.32 (d, *J* = 7.5 Hz, 1 H), 7.86 (d, *J* = 7.5 Hz, 1 H), 7.78 (s, 1 H), 7.58 (s, 1 H), 3.32 (t, *J* = 6.2 Hz, 2 H), 2.52 (t, *J* = 6.2 Hz, 2 H), 2.32 (s, 6 H), 2.15-2.13 (m, 2 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 163.2, 159.3, 148.2, 144.6, 131.2, 121.3, 114.6, 110.2, 106.0, 59.1, 47.0 (2 C), 29.4, 25.9
	MS	Found 291.24 $[M+H]^+$, Calculated for $C_{14}H_{17}N_3O_4$ 291.31 $[M+H]^+$
	Yield	196 mg (98% yield)

Synthesis of 5-nitro-*N*-(3-(pyrrolidin-1-yl)propyl)benzofuran-2-carboxamide (3.55)

Compound **3.55** was prepared using protocol described for the synthesis of **3.54**.

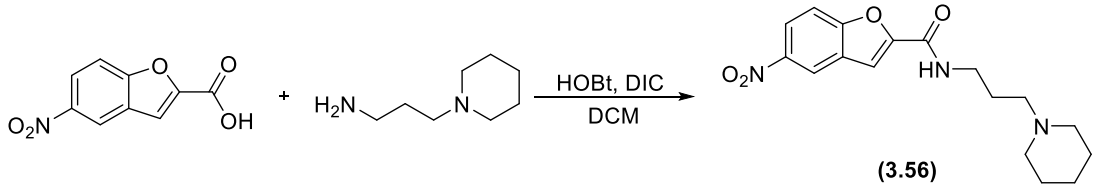


(3.55)

3.55 Amber oil	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 8.48 (s, 1 H), 8.24 (d, *J* = 7.4 Hz, 1 H), 7.86 (d, *J* = 7.5 Hz, 1 H), 7.78 (s, 1 H), 7.58 (s, 1 H), 3.12 (t, *J* = 6.1 Hz, 2 H), 2.54-2.52 (m, 4 H), 2.34 (t, *J* = 6.1 Hz, 2 H), 2.04-2.02 (m, 2 H), 1.76-1.72 (m, 4 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 166.2, 157.3, 150.2, 142.6, 129.2, 121.3, 118.6, 109.2, 105.8, 56.4 (2 C), 42.3, 28.4, 23.1 (2 C)
	MS	Found 317.19 $[M+H]^+$, Calculated for $C_{16}H_{19}N_3O_4$ 317.35 $[M+H]^+$
	Yield	215 mg (99% yield)

Synthesis of 5-nitro-*N*-(3-(piperidin-1-yl)propyl)benzofuran-2-carboxamide (3.56)

Compound **3.56** was prepared using protocol described for the synthesis of **3.54**.

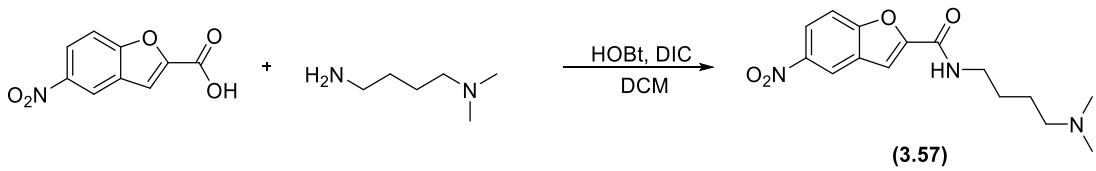


(3.56)

3.56 Yellow oil	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 8.46 (s, 1 H), 8.22 (d, *J* = 7.5 Hz, 1 H), 7.86 (d, *J* = 7.5 Hz, 1 H), 7.78 (s, 1 H), 7.58 (s, 1 H), 3.12 (t, *J* = 6.1 Hz, 2 H), 2.54-2.52 (m, 4 H), 2.34 (t, *J* = 6.1 Hz, 2 H), 2.36-2.32 (m, 2 H), 1.46-1.44 (m, 6 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 165.2, 159.3, 150.2, 147.6, 130.5, 119.3, 116.1, 110.2, 108.0, 57.1 (2 C), 45.3 (2 C), 28.4, 25.6 (2 C), 24.2
	MS	Found 331.19 $[M+H]^+$, Calculated for $C_{17}H_{21}N_3O_4$ 331.15 $[M+H]^+$
	Yield	226 mg (99% yield)

Synthesis of *N*-(4-(dimethylamino)butyl)-5-nitrobenzofuran-2-carboxamide (3.57)

Compound **3.57** was prepared using protocol described for the synthesis of **3.54**.

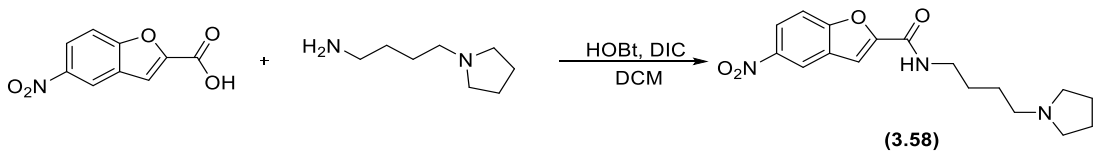


(3.57)

3.57 Pale yellow oil	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 8.77 (s, 1 H), 8.32 (d, *J* = 7.4 Hz, 1 H), 7.87 (d, *J* = 7.4 Hz, 1 H), 7.78 (s, 1 H), 7.58 (s, 1 H), 3.32 (t, *J* = 6.2 Hz, 2 H), 2.52 (t, *J* = 6.2 Hz, 2 H), 2.32 (s, 6 H), 1.78-1.74 (m, 2 H), 1.54-1.52 (m, 2 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 164.1, 159.7, 149.5, 144.2, 129.6, 122.1, 117.0, 109.8, 106.6, 61.2, 47.0, 32.1, 29.4, 25.9 (2 C)
	MS	Found 305.29 $[M+H]^+$, Calculated for $C_{15}H_{19}N_3O_4$ 305.33 $[M+H]^+$
	Yield	204 mg (97% yield)

Synthesis of 5-nitro-*N*-(4-(pyrrolidin-1-yl)butyl)benzofuran-2-carboxamide (3.58)

Compound **3.58** was prepared using protocol described for the synthesis of **3.54**.

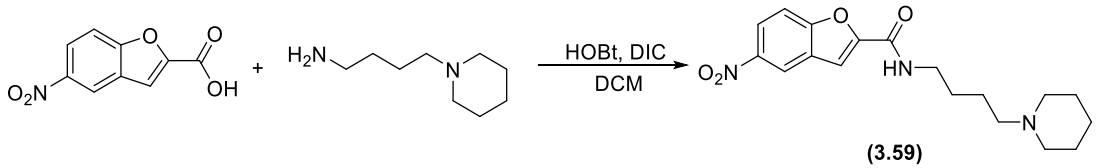


(3.58)

3.58 Yellow oil	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 8.48 (s, 1 H), 8.24 (d, *J* = 7.4 Hz, 1 H), 7.86 (d, *J* = 7.5 Hz, 1 H), 7.78 (s, 1 H), 7.58 (s, 1 H), 3.08 (t, *J* = 6.1 Hz, 2 H), 3.02 (t, *J* = 5.9 Hz, 2 H), 2.54-2.52 (m, 4 H), 1.64-1.62 (m, 2 H), 1.76-1.68 (m, 6 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 163.0, 158.1, 149.2, 145.6, 131.2, 122.1, 113.9, 109.6, 108.2, 56.8 (2 C), 55.2, 42.3, 27.3, 24.2, 23.5 (2 C)
	MS	Found 331.07 $[M+H]^+$, Calculated for $C_{17}H_{21}N_3O_4$ 331.15 $[M+H]^+$
	Yield	224 mg (98% yield)

Synthesis of 5-nitro-*N*-(4-(pyrrolidin-1-yl)butyl)benzofuran-2-carboxamide (3.59)

Compound **3.59** was prepared using protocol described for the synthesis of **3.54**.

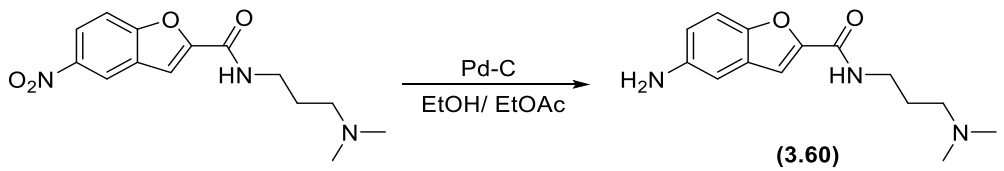


(3.59)

3.59 Pale yellow oil	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 8.46 (s, 1 H), 8.22 (d, *J* = 7.5 Hz, 1 H), 7.86 (d, *J* = 7.5 Hz, 1 H), 7.78 (s, 1 H), 7.58 (s, 1 H), 3.06 (t, *J* = 6.1 Hz, 2 H), 2.94 (t, *J* = 6.1 Hz, 2 H), 2.44-2.41 (m, 4 H), 2.36-2.32 (m, 2 H), 1.52-1.37 (m, 8 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 167.1, 159.6, 151.1, 146.5, 132.0, 119.3, 113.2, 110.9, 106.6, 56.8 (2 C), 55.0, 45.3, 29.4, 27.2, 25.2 (2 C), 23.9
	MS	Found 345.31 $[M+H]^+$, Calculated for $C_{18}H_{23}N_3O_4$ 345.40 $[M+H]^+$
	Yield	236 mg (99% yield)

Synthesis of 5-amino-*N*-(3-(dimethylamino)propyl)benzofuran-2-carboxamide (3.60)

For this second step, the monoamide intermediates **3.54** were hydrogenated in a parr hydrogenator reactor at 35 – 40 mbar until the H_2 uptake was finished (5 h). Up on completion, the reaction mixture was filtered through a layer of diatomaceous earth (Celite) and volatiles were removed under vacuum to give **3.60** as brown oil.

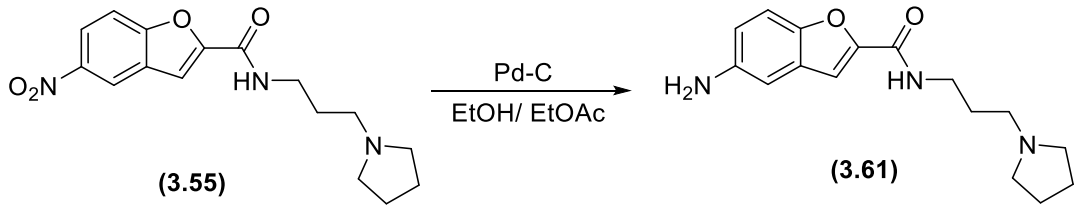


(3.60)

3.60 Brown oil	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 7.88 (s, 1 H), 7.86 (s, 1 H), 7.58 (s, 1 H), 7.48 (d, *J* = 7.5 Hz, 1 H), 6.58 (d, *J* = 7.5 Hz, 1 H), 4.59 (s, 2 H), 3.32 (t, *J* = 6.2 Hz, 2 H), 2.52 (t, *J* = 6.2 Hz, 2 H), 2.12-2.08 (m, 2 H), 2.12 (s, 6 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 151.2, 149.3, 148.2, 144.6, 131.2, 115.3, 110.6, 110.2, 106.0, 59.1, 47.0 (2 C), 29.4, 25.9
	MS	Found 261.12 $[M+H]^+$, Calculated for $C_{14}H_{19}N_3O_2$ 261.33 $[M+H]^+$
	Yield	174 mg (99% yield)

Synthesis of 5-amino-*N*-(3-(pyrrolidin-1-yl)propyl)benzofuran-2-carboxamide (3.61)

Hydrogenation protocol employed for the compound **3.60** was used.

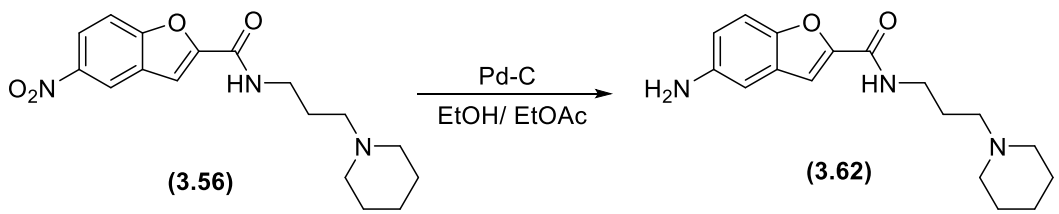


(3.55) **(3.61)**

3.61 Tan oil	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 7.81 (s, 1 H), 7.78 (s, 1 H), 7.58 (s, 1 H), 7.48 (d, *J* = 7.5 Hz, 1 H), 6.58 (d, *J* = 7.5 Hz, 1 H), 4.62 (s, 2 H), 3.12 (t, *J* = 6.1 Hz, 2 H), 2.54-2.52 (m, 4 H), 2.34 (t, *J* = 6.1 Hz, 2 H), 2.04-2.02 (m, 2 H), 1.76-1.72 (m, 4 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 158.6, 149.8, 146.2, 131.2, 116.1, 108.2, 58.1, 43.2, 30.5, 25.2, 24.9
	MS	Found 287.03 $[M+H]^+$, Calculated for $C_{16}H_{21}N_3O_2$ 287.36 $[M+H]^+$
	Yield	190 mg (97% yield)

Synthesis of 5-amino-*N*-(3-(piperidin-1-yl)propyl)benzofuran-2-carboxamide (3.62)

Hydrogenation protocol employed for the compound **3.60** was used.

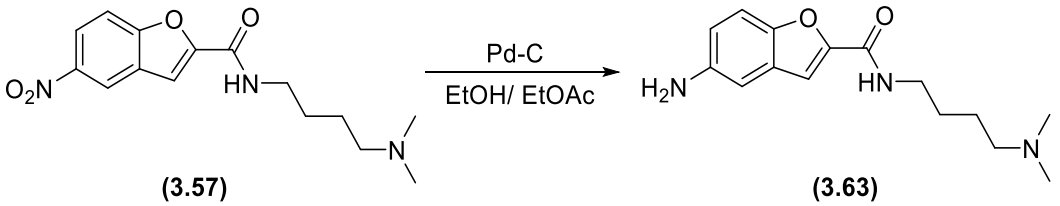


(3.56) **(3.62)**

3.62 Tan oil	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 7.83 (s, 1 H), 7.80 (s, 1 H), 7.62 (s, 1 H), 7.51 (d, *J* = 7.4 Hz, 1 H), 6.57 (d, *J* = 7.4 Hz, 1 H), 4.61 (s, 2 H), 3.12 (t, *J* = 6.2 Hz, 2 H), 2.56-2.53 (m 4 H), 2.45 (t, *J* = 6.2 Hz, 2 H), 2.36-2.32 (m, 2 H), 1.48-1.46 (m, 6 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 159.3, 150.2, 147.6, 143.4, 130.5, 119.3, 111.4, 108.7, 108.0, 57.1 (2 C), 45.3 (2 C), 28.4, 25.6 (2 C), 24.2
	MS	Found 301.22 $[M+H]^+$, Calculated for $C_{17}H_{23}N_3O_2$ 301.39 $[M+H]^+$
	Yield	200 mg (97% yield)

Synthesis of 5-amino-*N*-(4-(dimethylamino)butyl)benzofuran-2-carboxamide (3.63)

Hydrogenation protocol employed for the compound **3.60** was used.

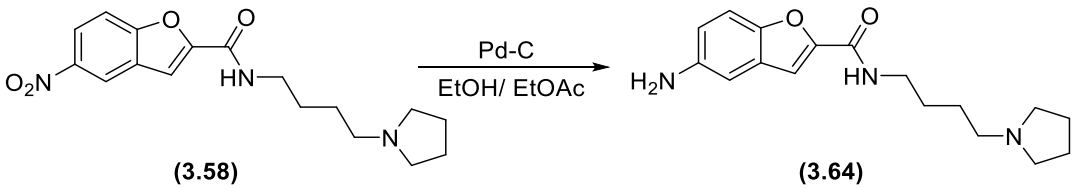


(3.57) **(3.63)**

3.63 Pale yellow oil	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 7.83 (s, 1 H), 7.76 (s, 1 H), 7.62 (s, 1 H), 7.49 (d, *J* = 7.4 Hz, 1 H), 6.56 (d, *J* = 7.4 Hz, 1 H), 4.63 (s, 2 H), 3.29 (t, *J* = 6.2 Hz, 2 H), 2.53 (t, *J* = 6.2 Hz, 2 H), 2.34 (s, 6 H), 1.78-1.76 (m, 2 H), 1.56-1.50 (m, 2 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 158.7, 151.2, 147.5, 144.2, 129.6, 122.1, 113.0, 110.8, 108.6, 61.2, 47.0, 32.1, 29.4, 26.0 (2 C)
	MS	Found 275.39 $[M+H]^+$, Calculated for $C_{15}H_{21}N_3O_2$ 275.45 $[M+H]^+$
	Yield	180 mg (98% yield)

Synthesis of 5-amino-*N*-(4-(pyrrolidin-1-yl)butyl)benzofuran-2-carboxamide (3.64)

Hydrogenation protocol employed for the compound **3.60** was used.

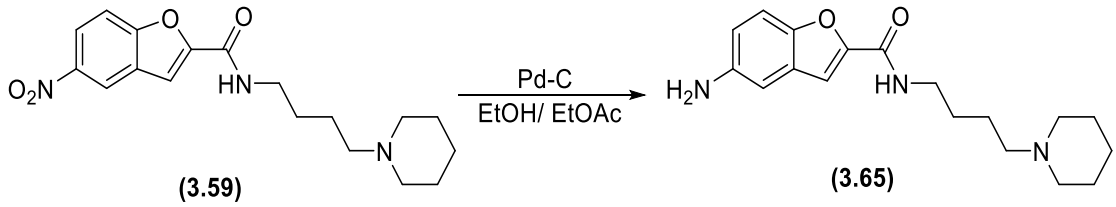


(3.58) **(3.64)**

3.64 Yellow oil	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 7.82 (s, 1 H), 7.76 (s, 1 H), 7.64 (s, 1 H), 7.50 (d, *J* = 7.4 Hz, 1 H), 6.66 (d, *J* = 7.4 Hz, 1 H), 4.62 (s, 2 H), 3.12 (t, *J* = 6.1 Hz, 2 H), 3.02 (t, *J* = 5.9 Hz, 2 H), 2.54-2.49 (m, 4 H), 1.64-1.60 (m, 2 H), 1.76-1.72 (m, 6 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 157.9, 152.0, 149.2, 145.6, 131.2, 122.1, 112.3, 108.6, 108.2, 56.8 (2 C), 55.2, 42.3, 27.3, 24.2, 23.5 (2 C)
	MS	Found 301.46 $[M+H]^+$, Calculated for $C_{17}H_{23}N_3O_2$ 301.39 $[M+H]^+$
	Yield	200 mg (98% yield)

Synthesis of 5-amino-*N*-(4-(dimethylamino)butyl)benzofuran-2-carboxamide (3.65)

Hydrogenation protocol employed for the compound **3.60** was used.

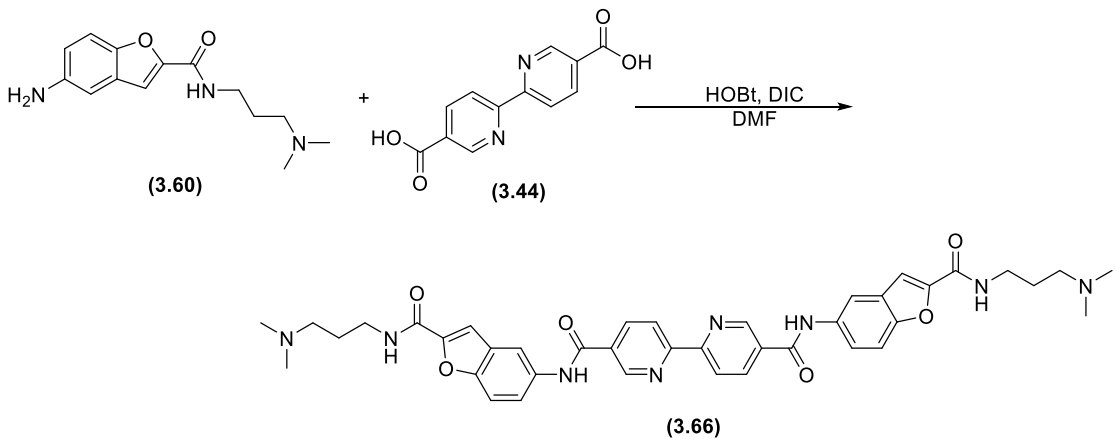


(3.59) **(3.65)**

3.65 Yellow oil	**^{1}H NMR**	(400 MHz, $CDCl_3$) δ 7.84 (s, 1 H), 7.80 (s, 1 H), 7.63 (s, 1 H), 7.51 (d, *J* = 7.3 Hz, 1 H), 6.59 (d, *J* = 7.3 Hz, 1 H), 4.60 (s, 2 H), 3.06 (t, *J* = 6.0 Hz, 2 H), 2.96 (t, *J* = 6.0 Hz, 2 H), 2.46-2.42 (m, 4 H), 2.36-2.32 (m, 2 H), 1.50-1.39 (m, 8 H)
	^{13}C NMR	(100 MHz, $CDCl_3$) δ 158.6, 151.4, 146.5, 142.0, 132.1, 119.3, 113.2, 108.9, 108.6, 56.8 (2 C), 55.0, 45.3, 29.4, 27.2, 25.2 (2 C), 23.9
	MS	Found 315.51 $[M+H]^+$, Calculated for $C_{18}H_{25}N_3O_2$ 315.42 $[M+H]^+$
	Yield	206 mg (96% yield)

Synthesis of *N*^{5},*N*$^{5'}$-*bis*(2-((3-(dimethylamino)propyl)carbamoyl)benzofuran-5-*yl*)-[2,2'-bipyridine]-*5,5'*-dicarboxamide (3.66)

To a stirred solution of [2,2'-bipyridine]-5,5'-dicarboxylic acid (30 mg, 0.12 mmol) in dimethylformamide (7 mL) was added, HOBt (67 mg, 0.49 mmol) and DIC (67 μL, 0.43 mmol) in one portion. After 30 minutes, **3.60** (65 mg, 0.25 mmol) was added and the resulting mixture was left to stir at room temperature for 13 h. The reaction was quenched with water (12 mL). Hexane (20 mL) was added to the reaction mixture and the resulting precipitate was recovered by filtration and washed with hexane (60 mL). The residue obtained was then dissolved in a mixture of acetonitrile/water (50/50) (40 mL), filtered to remove the impurities and concentrated to give the entitled product as a white solid.

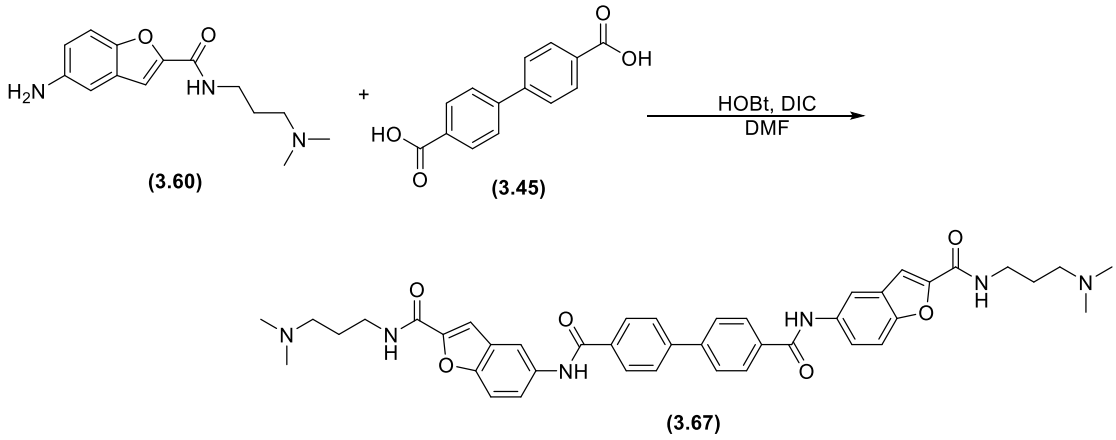


(3.60) (3.44) (3.66)

3.66 White solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 10.31 (s, 2 H), 8.92 (t, *J* = 5.6 Hz, 2 H), 8.62-8.56 (m, 2 H), 8.24 (d, *J* = 8.1 Hz, 4 H), 8.09 (d, *J* = 8.21 Hz, 4 H), 7.83 (s, 2 H), 7.66 (d, *J* = 7.5 Hz, 2 H), 3.28 (d, *J* = 5.3 Hz, 4 H), 2.48-2.46 (m, 4 H), 2.28 (s, 12 H), 1.69-1.66 (m, 4 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 165.6, 159.9, 152.2, 150.2, 144.6 (2 C), 132.8, 131.1, 130.1, 128.0, 117.9 (2 C), 114.2, 112.0, 109.4, 59.2, 48.2, 40.7 (2 C), 27.3
	IR	(υ_{max}/cm^{-1}) 3332, 3040, 2878, 1643, 1589, 1543, 1469, 1435, 1354
	MS	Found 731.10 $[M+H]^+$, Calculated for $C_{40}H_{42}N_8O_6$ 730.83 $[M+H]^+$
	HRMS	Observed 731.3287 $[M+H]^+$, Theoretical 731.3300 $[C_{40}H_{42}N_8O_6+H]^+$
	Yield	63 mg (51% yield)

Synthesis of N^4,N$^{4'}$-bis(2-((3-(dimethylamino)propyl)carbamoyl)benzofuran-5-yl)-[1,1'-biphenyl]-4,4'-dicarboxamide (3.67)

Compound **3.67** was prepared using protocol described for the synthesis of **3.66**.

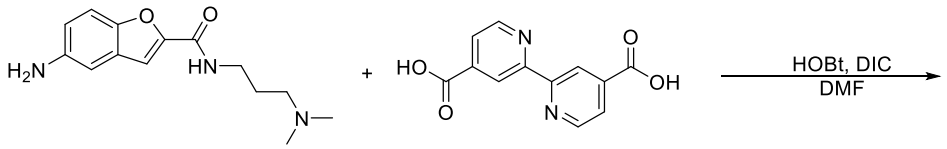


(3.60) (3.45) (3.67)

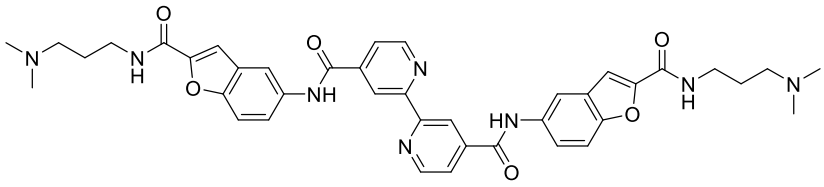
3.67 Tan solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 10.15 (s, 2 H), 8.62 (t, *J* = 5.6 Hz, 2 H), 8.58-8.52 (m, 2 H), 8.32 (d, *J* = 8.1 Hz, 4 H) 8.19 (d, *J* = 8.2 Hz, 4 H), 7.78-7.74 (m, 2H), 7.68 (s, 2 H), 7.54 (d, *J* = 7.5 Hz, 2 H), 3.23 (d, *J* = 5.3 Hz, 4 H), 2.36-2.32 (m, 4 H), 2.21 (s, 12 H), 1.67-1.63 (m, 4 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 164.9, 158.9, 152.3, 150.0, 143.7, 132.2, 131.0, 130.1 (2 C), 128.5 (2 C), 117.9 (2 C), 115.2, 112.3, 109.6, 59.5, 47.7 (2 C), 40.4, 26.3
	IR	(υ_{max}/cm^{-1}) 3247, 3073, 2858, 1654, 1546, 1469, 1363
	MS	Found 729.50 $[M-H]^+$, Calculated for $C_{42}H_{44}N_6O_6$ 729.85 $[M+H]^+$
	HRMS	Observed 729.3382 $[M+H]^+$, Theoretical 729.3395 $[C_{42}H_{44}N_6O_6+H]^+$
	Yield	91 mg (76.4% yield)

Synthesis of N^4,N$^{4'}$-bis(2-((3-(dimethylamino)propyl)carbamoyl)benzofuran-5-yl)-[2,2'-bipyridine]-4,4'-dicarboxamide (3.68)

Compound **3.68** was prepared using protocol described for the synthesis of **3.66**.



(3.60) **(3.45)**

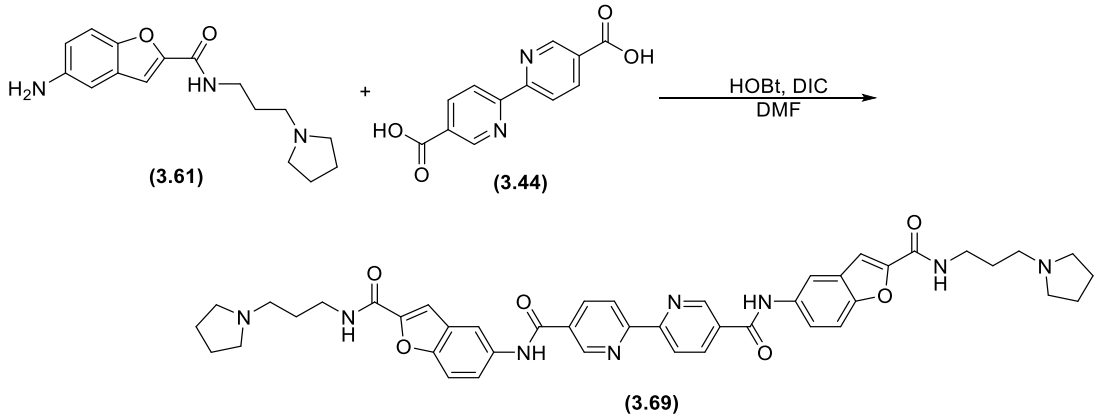


(3.68)

3.68 Pale pink solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 10.20 (s, 2 H), 8.87 (d, *J* = 5.6 Hz, 2 H), 8.62-8.56 (m, 2 H), 8.64 (d, *J* = 8.1 Hz, 4 H), 7.80 (d, *J* = 8.2 Hz, 2 H), 7.79 (s, 2 H), 7.74 (d, *J* = 7.5 Hz, 2H), 7.59 (d, *J* = 7.5 Hz, 2 H), 3.28 (d, *J* = 5.3 Hz, 4 H), 2.46 (m, 4 H), 2.28 (s, 12 H), 1.69-1.66 (m, 4 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 166.6, 155.8, 150.0, 149.8, 145.8, 132.7, 131.1, 122.1 (2 C), 118.2, 117.0 (2 C), 111.2, 112.0, 109.4, 58.1, 47.2 (2 C), 40.3, 26.2
	IR	(υ_{max}/cm^{-1}) 3255, 3063, 2878, 1655, 1547, 1469, 1362
	MS	Found 729.5 $[M-H]^+$, Calculated for $C_{40}H_{42}N_8O_6$ 730.83 $[M+H]^+$
	HRMS	Observed 731.3296 $[M+H]^+$, Theoretical 731.3300 $[C_{40}H_{42}N_8O_6+H]^+$
	Yield	39 mg (44% yield)

Synthesis of *N*^{5},*N*$^{5'}$-*bis*(2-((3-(pyrrolidin-1-yl)propyl)carbamoyl)benzofuran-5-yl)-[2,2'-bipyridine]-5,5'-dicarboxamide (3.69)

Compound **3.69** was prepared using protocol described for the synthesis of **3.66**.

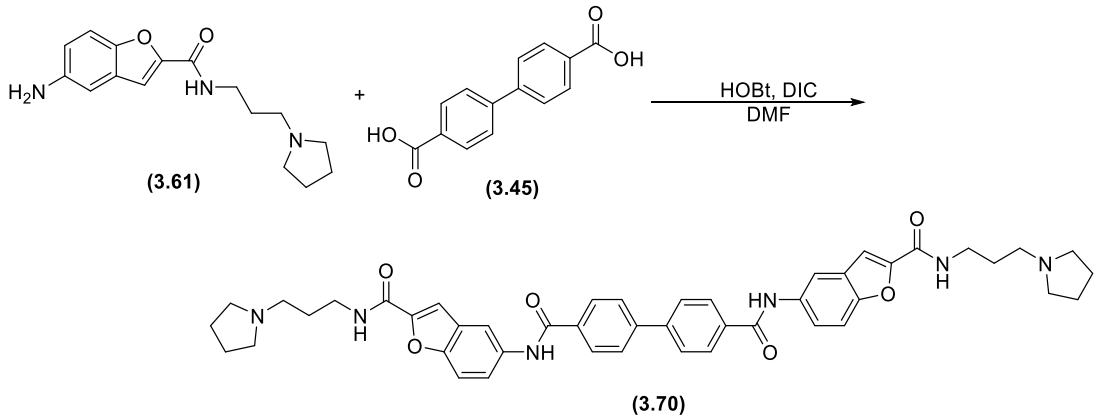


(3.61) (3.44) (3.69)

3.69 Pale yellow solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 10.21 (s, 2 H), 8.78 (t, *J* = 5.6 Hz, 2 H), 8.32-8.26 (m, 2 H), 8.14 (d, *J* = 8.2 Hz, 4 H), 8.11 (d, *J* = 8.2 Hz, 4 H), 7.69 (s, 2 H), 7.57 (d, *J* = 7.6 Hz, 2 H), 3.09 (d, *J* = 5.3 Hz, 4 H), 2.66-2.61 (m, 4 H), 2.48-2.44 (m, 8 H), 1.78-1.74 (m, 4 H), 1.71-1.69 (m, 8 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 168.6, 159.9, 155.3, 154.2, 139.6 (2 C), 135.4, 132.8, 132.0, 128.0, 125.1, 116.9 (2 C), 113.2, 109.5, 58.0 (2 C), 39.8 (2 C), 27.2, 23.4 (2 C)
	IR	(υ_{max}/cm^{-1}) 3317, 3086, 2974, 2881, 1643, 1589, 1535, 1469, 1350
	MS	Found 783.20 $[M+H]^+$, Calculated for $C_{44}H_{46}N_8O_6$ 783.90 $[M+H]^+$
	HRMS	Observed 783.3600 $[M+H]^+$, Theoretical 783.3613 $[C_{44}H_{46}N_8O_6+H]^+$
	Yield	79 mg (77% yield)

Synthesis of *N*^{4},*N*$^{4'}$-*bis*(2-((3-(pyrrolidin-1-yl)propyl)carbamoyl)benzofuran-5-yl)-[1,1'-biphenyl]-4,4'-dicarboxamide (3.70)

Compound **3.70** was prepared using protocol described for the synthesis of **3.66**.

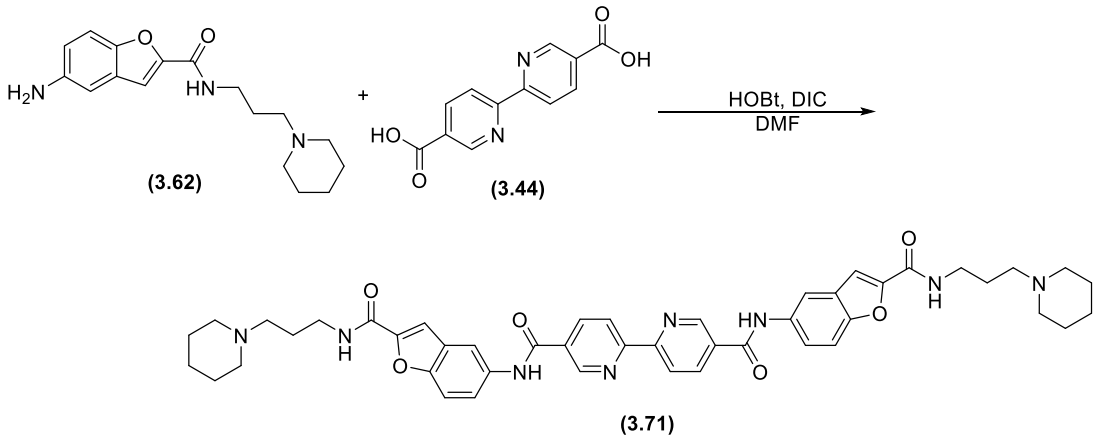


(3.61) (3.45) (3.70)

3.70 Tan solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 10.34 (s, 2 H), 8.76 (t, *J* = 5.5 Hz, 2 H), 8.32-8.27 (m, 2 H), 8.22 (d, *J* = 8.2 Hz, 4 H), 8.01 (d, *J* = 8.2 Hz, 2 H), 7.79 (s, 2 H), 7.66 (d, *J* = 5.4 Hz, 4 H), 7.57 (d, *J* = 7.5 Hz, 2 H), 3.19 (d, *J* = 5.2 Hz, 4 H), 2.86-2.82 (m, 4 H), 2.59-2.55 (m, 8 H), 1.86-1.82 (m, 4 H), 1.73-1.69 (m, 8 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 165.3, 158.2, 152.8, 150.0, 144.5, 134.2, 132.5, 131.2 (2 C), 128.0 (2 C), 118.4 (2 C), 115.5, 111.5, 109.4, 58.3 (2 C), 40.0 (2 C), 27.6, 23.6 (2 C)
	IR	(υ_{max}/cm^{-1}) 3290, 3036, 2970, 2881, 1647, 1535, 1469, 1365
	MS	Found 781.20 $[M+H]^+$, Calculated for $C_{46}H_{48}N_6O_6$ 781.93 $[M+H]^+$
	HRMS	Observed 781.3698 $[M+H]^+$, Theoretical 781.3708 $[C_{46}H_{48}N_6O_6+H]^+$
	Yield	69 mg (74% yield)

Synthesis of *N^5,N$^{5'}$-bis*(2-((3-(piperidin-1-yl)propyl)carbamoyl)benzofuran-5-*yl*)-[2,2'-bipyridine]-5,5'-dicarboxamide (3.71)

Compound **3.71** was prepared using protocol described for the synthesis of **3.66**.

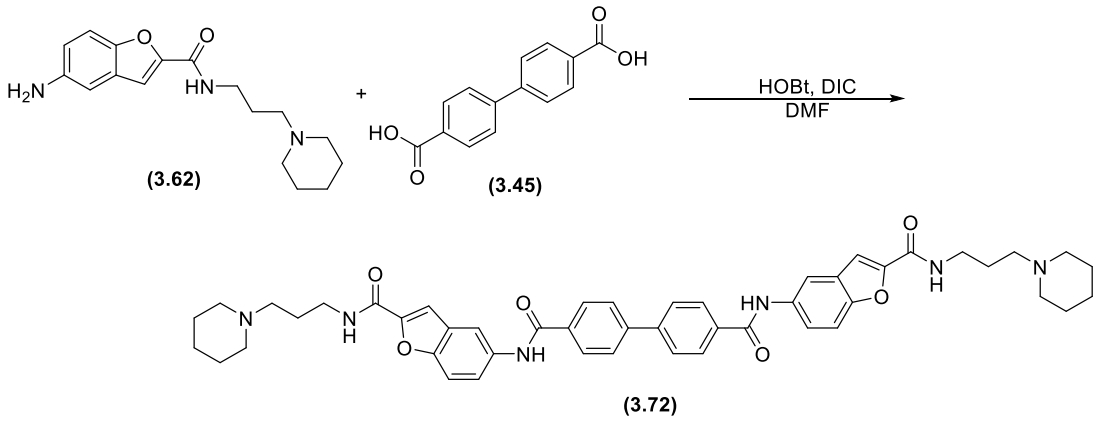


(3.62) (3.44) (3.71)

3.71 White powder	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 10.49 (s, 2 H), 8.80 (t, *J* = 5.5 Hz, 2 H), 8.34-8.28 (m, 2 H), 8.13 (d, *J* = 8.3 Hz, 4 H), 8.01 (d, *J* = 8.3 Hz, 4 H), 7.69 (s, 2 H), 7.57 (d, *J* = 7.5 Hz, 2 H), 3.49-3.45 (m, 10 H), 2.51-2.46 (m, 4 H), 1.90-1.82 (m, 4 H), 1.69-1.62 (m, 4 H), 1.52 (d, *J* = 4.8 Hz, 10 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 166.5, 158.8, 154.2, 150.6 (2 C), 139.5, 132.8, 131.2, 128.0, 126.9, 117.1 (2 C), 115.3, 112.1, 108.5, 57.6 (2 C), 39.8 (2 C), 26.6, 24.5, 23.3 (2 C)
	IR	(υ_{max}/cm^{-1}) 3231, 3038, 2943, 2861, 1644, 1539, 1472, 1303
	MS	Found 811.40 $[M+H]^+$, Calculated for $C_{46}H_{50}N_8O_6$ 811.96 $[M+H]^+$
	HRMS	Observed 811.3925 $[M+H]^+$, Theoretical 811.3926 $[C_{46}H_{50}N_8O_6+H]^+$
	Yield	29 mg (34% yield)

Synthesis of *N^4,N$^{4'}$-bis*(2-((3-(piperidin-1-yl)propyl)carbamoyl)benzofuran-5-yl)-[1,1'-biphenyl]-4,4'-dicarboxamide (3.72)

Compound **3.72** was prepared using protocol described for the synthesis of **3.66**.

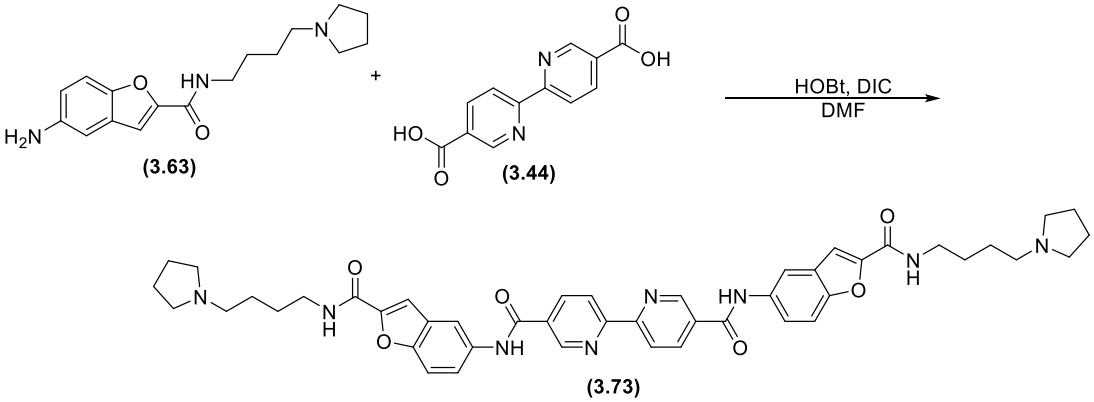


(3.62) **(3.45)** **(3.72)**

3.72 Tan solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 10.53 (s, 2 H), 8.84 (t, *J* = 5.4 Hz, 2 H), 8.38-8.32 (m, 2 H), 8.23 (d, *J* = 8.3 Hz, 4 H), 8.06 (d, *J* = 8.3 Hz, 4 H), 7.78 (s, 2 H), 7.58 (dd, *J* = 7.5 Hz, 4 H), 3.53-3.49 (m, 10 H), 2.52-2.47 (m, 4 H), 1.93-1.85 (m, 4 H), 1.68-1.61 (m, 4 H), 1.54-1.50 (m, 10 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 169.5, 157.8, 154.2, 150.6, 139.6, 132.8, 130.2, 128.0 (2 C), 126.9 (2 C), 117.4 (2 C), 112.6, 108.6, 58.2 (2 C), 39.4 (2 C), 39.5, 27.7, 25.4 (2 C), 24.6
	IR	(υ_{max}/cm^{-1}) 3234, 3041, 2948, 2864, 1649, 1549, 1470, 1301
	MS	Found 809.51 $[M+H]^+$, Calculated for $C_{48}H_{52}N_6O_6$ 809.98 $[M+H]^+$
	HRMS	Observed 809.4027 $[M+H]^+$, Theoretical 809.4021 $[C_{48}H_{52}N_6O_6+H]^+$
	Yield	39 mg (42% yield)

Synthesis of *N*^{5},*N*$^{5'}$-*bis*(2-((4-(pyrrolidin-1-yl)butyl)carbamoyl)benzofuran-5-yl)-[2,2'-bipyridine]-5,5'-dicarboxamide (3.73)

Compound **3.73** was prepared using protocol described for the synthesis of **3.66**.

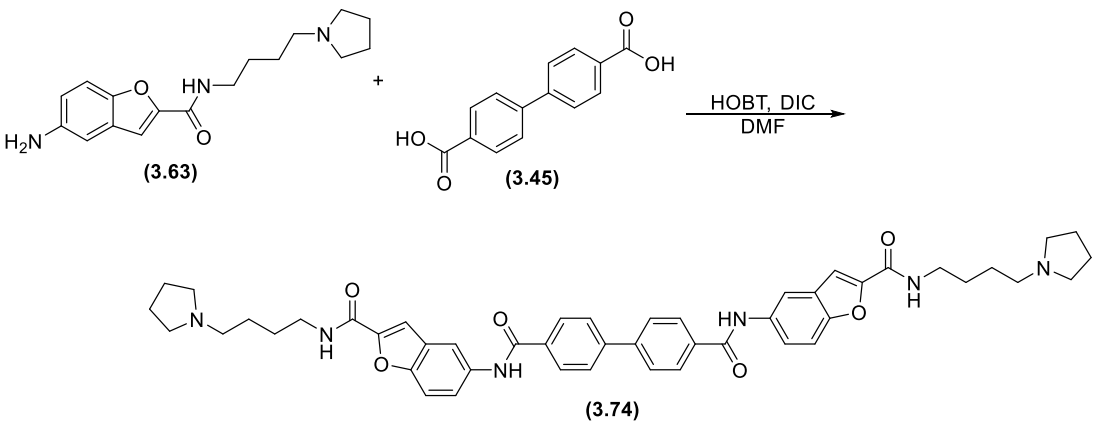


(3.63) (3.44) (3.73)

3.73 White powder	**H^{1} NMR**	(400 MHz, DMSO-d$_6$) δ 10.32 (s, 2 H), 8.80 (t, *J* = 5.5 Hz, 2 H), 8.34-8.28 (m, 2 H), 8.13 (d, *J* = 8.3 Hz, 4 H), 8.01 (d, *J* = 8.3 Hz, 4 H), 7.69 (s, 2 H), 7.57 (d, *J* = 7.5 Hz, 2 H), 3.09 (d, *J* = 5.3 Hz, 4 H), 2.96-2.94 (m, 8 H), 2.88-2.84 (m, 4 H), 1.78-1.72 (m, 8 H), 1.69-1.66 (m, 8 H)
	^{13}C NMR	(100 MHz, DMSO-d$_6$) δ 168.2, 159.6, 155.3, 154.1 (2 C), 139.5, 133.9, 132.5, 131.7, 128.0, 124.9, 118.5 (2 C), 112.1, 109.7, 58.6 (2 C), 56.3, 39.8, 28.6, 27.2, 23.3 (2 C)
	IR	(υ_{max}/cm^{-1}) 3317, 3086, 2947, 2874, 1643, 1589, 1535, 1469, 1350
	MS	Found 811.22 $[M+H]^+$, Calculated for $C_{46}H_{50}N_8O_6$ 811.96 $[M+H]^+$
	HRMS	Observed 811.3909 $[M+H]^+$, Theoretical 811.3926 $[C_{46}H_{50}N_8O_6+H]^+$
	Yield	127 mg (96% yield)

Synthesis of *N*^{4},*N*$^{4'}$-*bis*(2-((4-(pyrrolidin-1-yl)butyl)carbamoyl)benzofuran-5-yl)-[1,1'-biphenyl]-4,4'-dicarboxamide (3.74)

Compound **3.74** was prepared using protocol described for the synthesis of **3.66**.

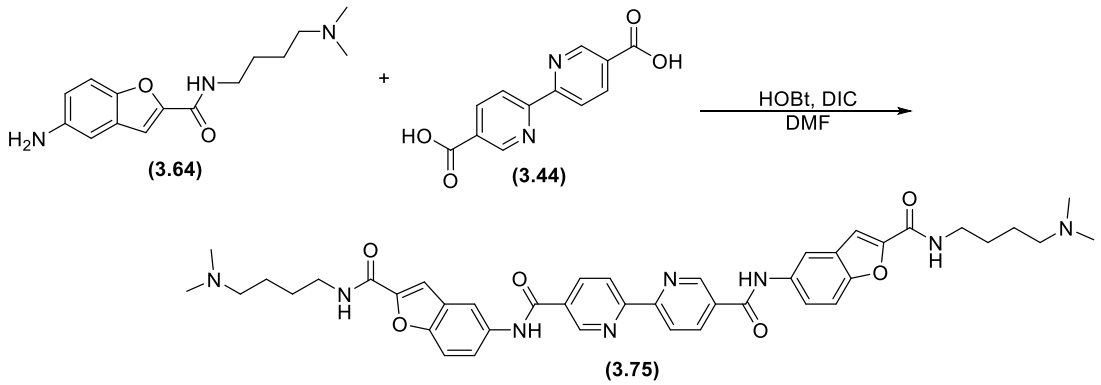


(3.63) (3.45) (3.74)

3.74 Pale yellow solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 10.25 (s, 2 H), 8.80 (t, *J* = 5.5 Hz, 2 H), 8.36-8.30 (m, 2 H), 8.13 (d, *J* = 8.3 Hz, 4 H), 7.96 (d, *J* = 8.3 Hz, 4 H), 7.65 (d, *J* = 9.1 Hz, 2 H), 7.56 (s, 2 H), 7.47 (d, *J* = 7.5 Hz, 2 H), 3.29 (d, *J* = 5.2 Hz, 4 H), 2.92-2.88 (m, 8 H), 2.84-2.80 (m, 4 H), 1.80-1.76 (m, 8 H), 1.59-1.55 (m, 8 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 173.0, 155.8, 149.5, 145.5, 145.2, 134.7, 128.0 (2 C), 124.9 (2 C), 118.0 (2 C), 116.1, 110.1, 109.2, 108.7, 59.3 (2 C), 55.0, 48.2, 38.7, 28.2, 22.4 (2 C)
	IR	(υ_{max}/cm^{-1}) 3255, 3039, 2947, 2874, 1647, 1539, 1469, 1308
	MS	Found 809.31 $[M+H]^+$, Calculated for $C_{48}H_{52}N_6O_6$ 809.98 $[M+H]^+$
	HRMS	Observed 809.3999 $[M+H]^+$, Theoretical 809.4021 $[C_{48}H_{52}N_6O_6+H]^+$
	Yield	86 mg (81% yield)

Synthesis of *N5,N5'-bis*(2-((4-(dimethylamino)butyl)carbamoyl)benzofuran-5-yl)-[2,2'-bipyridine]-5,5'-dicarboxamide (3.75)

Compound **3.75** was prepared using protocol described for the synthesis of **3.66**.

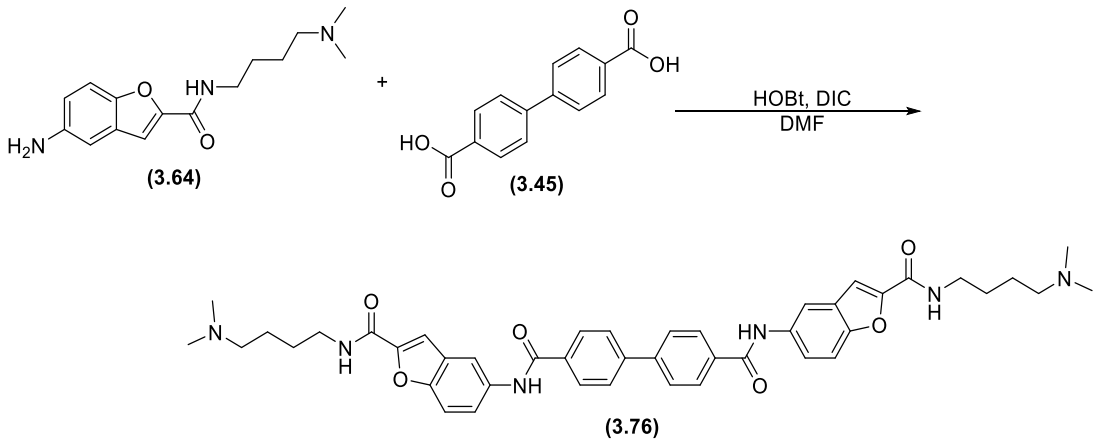


(3.64) **(3.44)** **(3.75)**

3.75 Pale yellow solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 10.19 (s, 2 H), 8.32-8.28 (m, 2 H), 8.01 (d, *J* = 8.1 Hz, 2 H), 7.98 (d, *J* = 7.7 Hz, 2 H), 7.88-7.82 (m, 2 H), 7.82 (s, 2 H), 7.57 (d, *J* = 7.5 Hz, 2 H), 3.01-2.98 (m, 8 H), 2.77-2.75 (m, 4 H), 2.09 (s, 12 H), 1.78-1.74 (m, 4 H), 1.59-1.55 (m, 4 H)
	^{13}C NMR	(100 MHz, DMSO-d6) δ 164.2, 160.1, 158.8, 155.4, 153.1, 150.6, 138.7, 133.4, 131.7, 127.2, 122.9, 119.1 (2 C), 113.0, 109.5, 62.2, 47.4 (2 C), 39.9, 27.5, 25.4
	IR	(υ_{max}/cm^{-1}) 3321, 3086, 2939, 1643, 1589, 1531, 1469, 1307
	MS	Found 759.32 $[M+H]^+$, Calculated for $C_{42}H_{46}N_8O_6$ 759.88 $[M+H]^+$
	HRMS	Observed 759.3600 $[M+H]^+$, Theoretical 759.3613 $[C_{42}H_{46}N_8O_6+H]^+$
	Yield	42 mg (54% yield)

Synthesis of *N*^{4},*N*$^{4'}$-*bis*(2-((4-(dimethylamino)butyl)carbamoyl)benzofuran-5-yl)-[1,1'-biphenyl]-4,4'-dicarboxamide (3.76)

Compound **3.76** was prepared using protocol described for the synthesis of **3.66**.

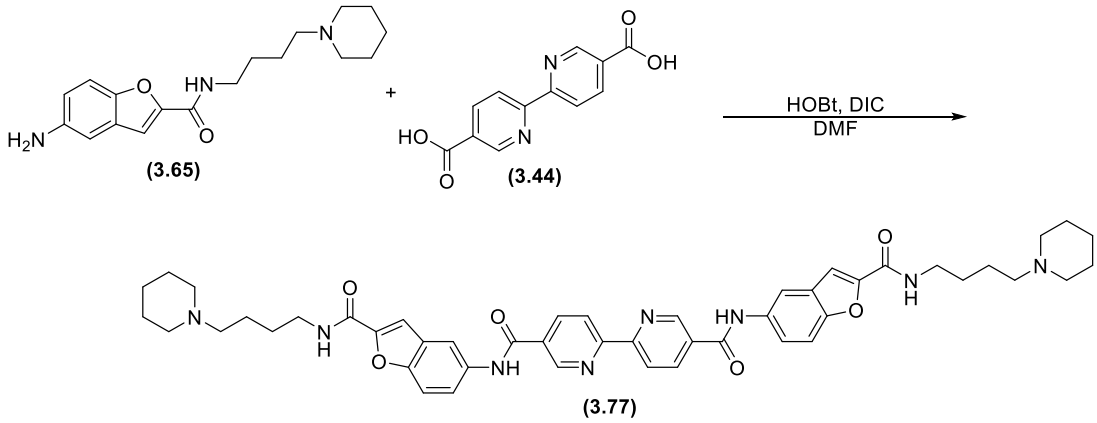


(3.64) (3.45) (3.76)

3.76 Grey solid	**^{1}H NMR**	(400 MHz, DMSO-d6) δ 10.23 (s, 2 H), 8.78-8.72 (m, 2 H), 8.01 (d, *J* = 8.3 Hz, 4 H), 7.98 (d, *J* = 7.7 Hz, 2 H), 7.92 (d, *J* = 7.2 Hz, 2 H), 7.90 (d, *J* = 7.1 Hz, 4 H), 7.82 (d, *J* = 7.4 Hz, 2 H), 7.59 (s, 2 H), 3.01-2.99 (m, 4 H), 2.62-2.58 (m, 4 H), 2.19 (s, 12 H), 1.80-1.76 (m, 4 H), 1.56-1.51 (m, 4 H)
	^{13}C NMR	(100 MHz, DMSO-d6) δ 163.2, 158.8, 152.1, 150.0, 144.5, 133.6, 131.4, 130.7 (2 C), 128.2 (2 C), 117.9 (2 C), 115.2, 112.0, 109.5, 62.6, 48.3, 39.1, 28.1, 24.3 (2 C)
	IR	(υ_{max}/cm^{-1}) 3325, 2939, 2758, 1647, 1589, 1535, 1469, 1312
	MS	Found 757.41 $[M+H]^+$, Calculated for $C_{44}H_{48}N_6O_6$ 756.90 $[M+H]^+$
	HRMS	Observed 757.3690 $[M+H]^+$, Theoretical 757.3708 $[C_{44}H_{48}N_6O_6+H]^+$
	Yield	67 mg (65% yield)

Synthesis of *N*^{5},*N*$^{5'}$-*bis*(2-((4-(piperidin-1-yl)butyl)carbamoyl)benzofuran-5-yl)-[2,2'-bipyridine]-5,5'-dicarboxamide (3.77)

Compound **3.77** was prepared using protocol described for the synthesis of **3.66**.

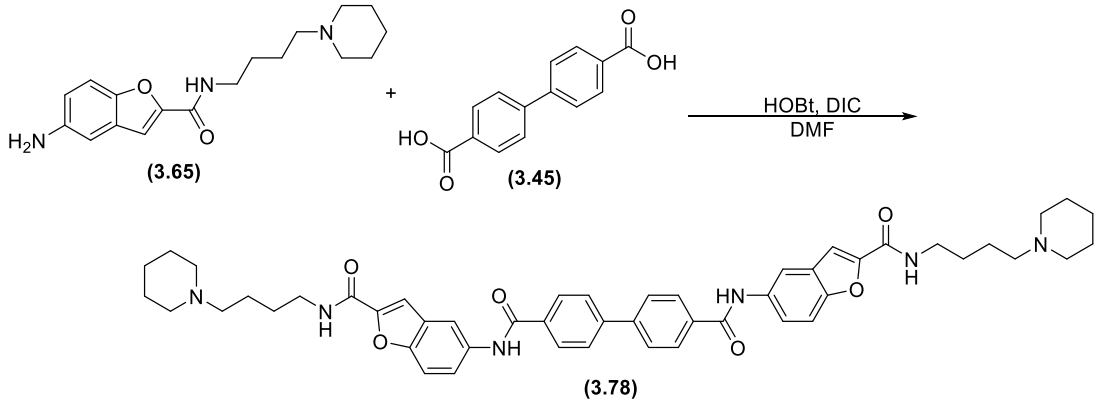


(3.65) (3.44) (3.77)

3.77 Pale yellow solid	**^{1}H NMR**	(400 MHz, DMSO-d6) δ 10.26 (s, 2 H), 8.86 (d, *J* = 8.3 Hz, 4 H), 8.16 (dd, *J* = 8.1 Hz, 4 H), 7.78 (s, 2 H), 7.68 (dd, *J* = 7.5 Hz, 4 H), 7.62 (s, 2 H), 3.12 (t, *J* = 6.8 Hz, 4 H), 3.09 (t, *J* = 6.4 Hz, 4 H), 2.56-2.50 (m, 4 H), 2.48-2.40 (m, 8 H), 1.88 (d, *J* = 5.6 Hz, 4 H), 1.82 (d, *J* = 4.6 Hz, 4 H), 1.59-1.53 (m, 2 H) 1.52-1.48 (m, 8 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 165.2, 159.1, 155.2, 150.0 (2 C), 138.4, 132.2, 130.6 (2 C), 123.4, 117.3 (2 C), 113.8, 112.4, 109.3, 58.9, 58.6, 39.6, 28.4 (2 C), 26.1, 25.6 (2 C), 24.2
	IR	(υ_{max}/cm^{-1}) 3327, 2901, 1640, 1591, 1539, 1489, 1309
	MS	Found 839.20 $[M+H]^+$, Calculated for $C_{48}H_{54}N_8O_6$ 839.01 $[M+H]^+$
	HRMS	Observed 839.4230 $[M+H]^+$, Theoretical 839.4239 $[C_{48}H_{54}N_8O_6+H]^+$
	Yield	79 mg (69% yield)

Synthesis of *N^4,N$^{4'}$-bis*(2-((4-(piperidin-1-yl)butyl)carbamoyl)benzofuran-5-yl)-[1,1'-biphenyl]-4,4'-dicarboxamide (3.78)

Compound **3.78** was prepared using protocol described for the synthesis of **3.66**.



(3.65) (3.45) (3.78)

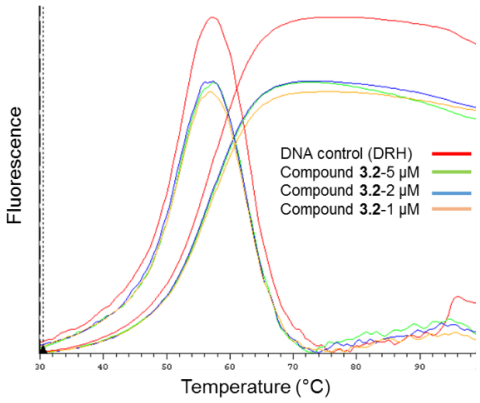
3.78 Pale yellow solid	**^{1}H NMR**	(400 MHz, DMSO-d_6) δ 10.42 (s, 2 H), 8.38-8.32 (m, 2 H), 8.23 (d, *J* = 8.1 Hz, 4 H), 8.06 (dd, *J* = 8.3 Hz, 4 H), 7.78 (s, 2 H), 7.68 (dd, *J* = 7.5 Hz, 4 H), 3.12 (t, *J* = 6.8 Hz, 4 H), 3.03 (t, *J* = 6.4 Hz, 4 H), 2.52-2.47 (m, 4 H), 2.44-2.40 (m, 8 H), 1.82 (d, *J* = 5.5 Hz, 4 H), 1.78 (d, *J* = 4.6 Hz, 4 H), 1.64-1.58 (m, 2 H), 1.54-1.50 (m, 8 H)
	^{13}C NMR	(100 MHz, DMSO-d_6) δ 165.6, 158.1, 152.2, 150.2, 144.6, 134.2, 131.8, 130.6 (2 C), 128.3 (2 C), 118.5 (2 C), 114.6, 112.8, 109.3, 59.0 (2 C), 58.3, 39.5, 28.7, 26.3, 25.4 (2 C), 24.1
	IR	(υ_{max}/cm^{-1}) 3329, 2932, 1647, 1593, 1535, 1469, 1303
	MS	Found 837.32 $[M+H]^+$, Calculated for $C_{50}H_{56}N_6O_6$ 837.03 $[M+H]^+$
	HRMS	Observed 837.4313 $[M+H]^+$, Theoretical 837.4334 $[C_{50}H_{56}N_6O_6+H]^+$
	Yield	47 mg (63% yield)

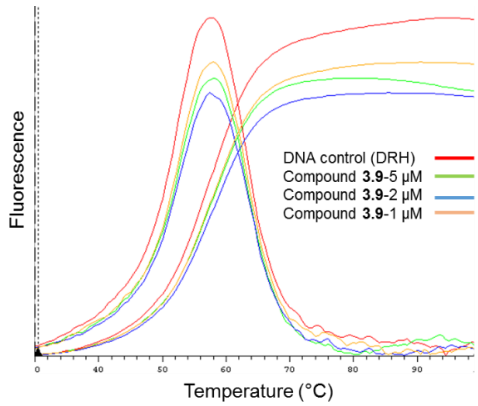
Appendix 2: Experimental data and Graphs

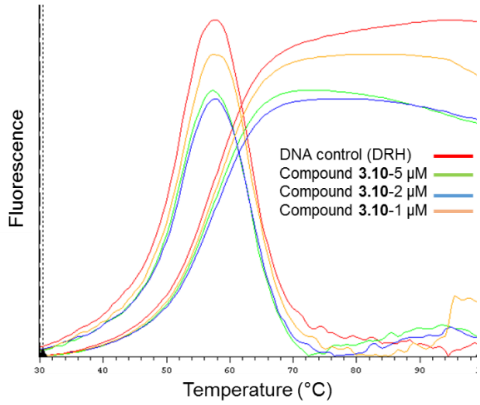
A2.1: FRET Melting Result

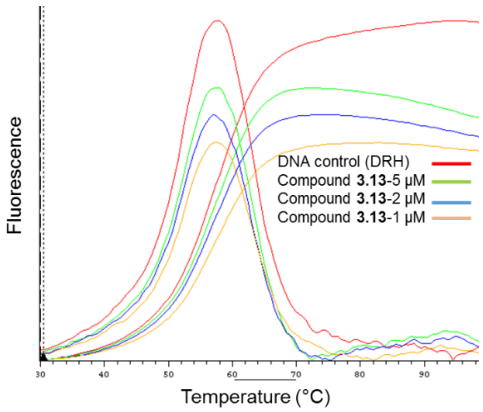
FRET melting curves are given below for the synthesized compounds. To be noted that all curves has got two different line diagrams in a single graph, *i.e.*, sigmoidal and parabolic shape. Sigmoidal shaped curve represents the FRET melting and parabolic shaped curve represents the change of fluorescence intensity over the change of temperature (-dI/dT).

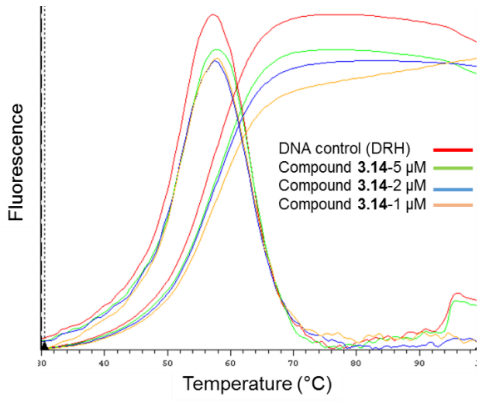
A2.1.1: DNA/RNA Hybrid duplex

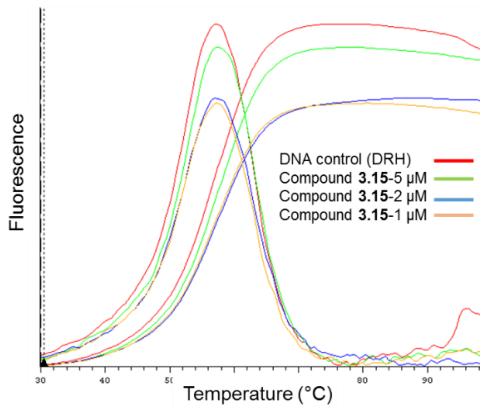


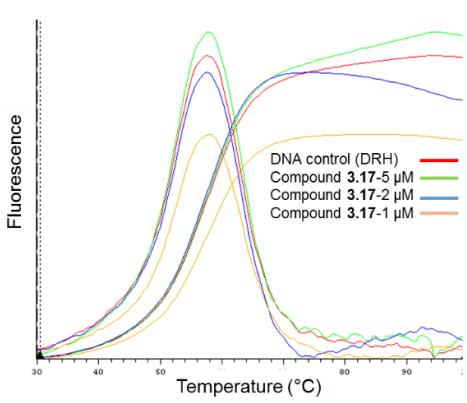


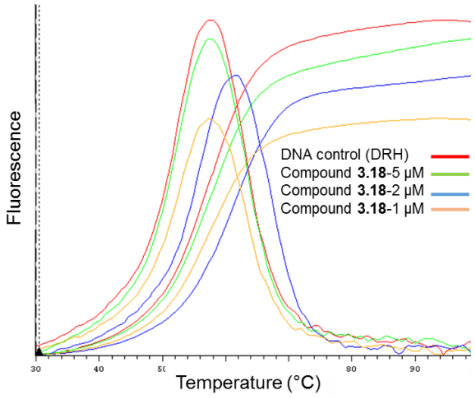


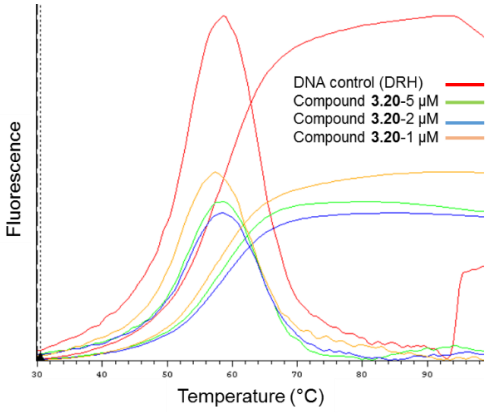


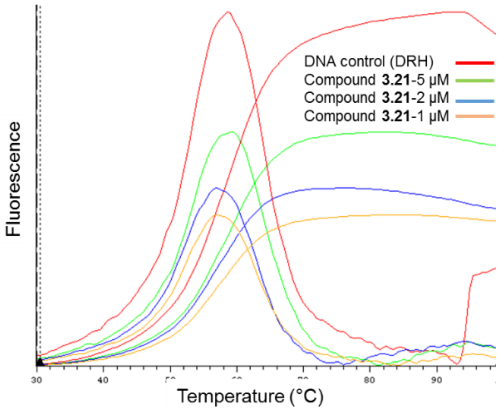


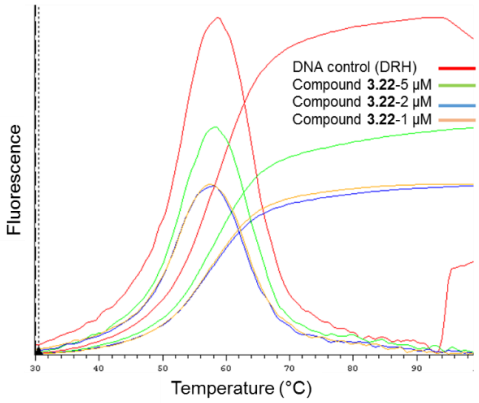


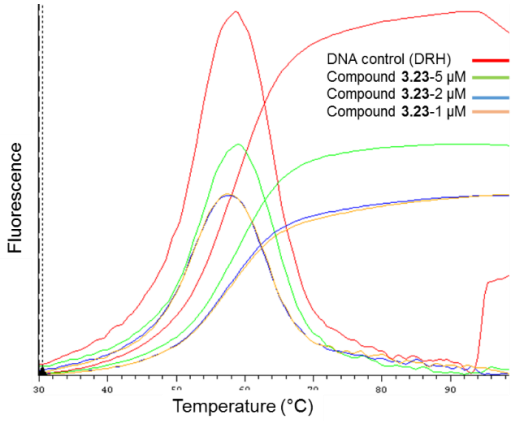

Fluorescence
Temperature (°C)
DNA control (DRH)
Compound 3.17-5 µM
Compound 3.17-2 µM
Compound 3.17-1 µM

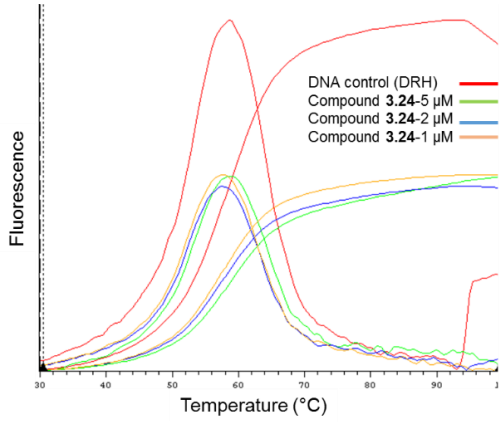

Fluorescence
Temperature (°C)
DNA control (DRH)
Compound 3.18-5 µM
Compound 3.18-2 µM
Compound 3.18-1 µM

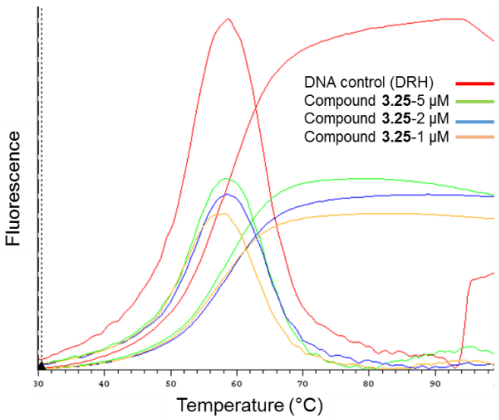

Fluorescence
Temperature (°C)
DNA control (DRH)
Compound 3.20-5 µM
Compound 3.20-2 µM
Compound 3.20-1 µM

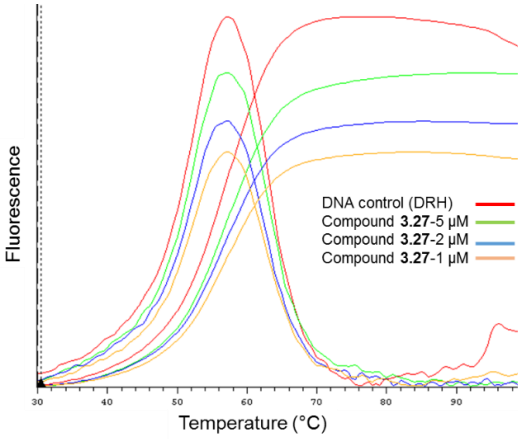

Fluorescence
Temperature (°C)
DNA control (DRH)
Compound 3.21-5 µM
Compound 3.21-2 µM
Compound 3.21-1 µM

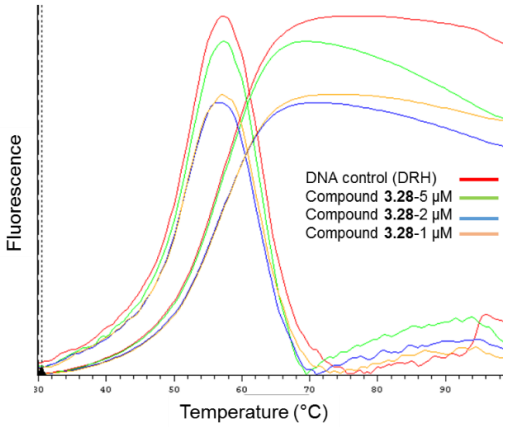

Fluorescence
Temperature (°C)
DNA control (DRH)
Compound 3.22-5 µM
Compound 3.22-2 µM
Compound 3.22-1 µM

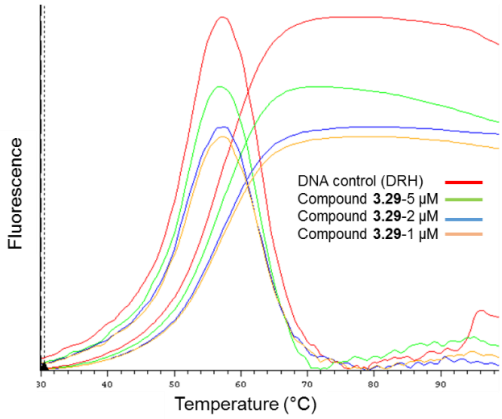

Fluorescence
Temperature (°C)
DNA control (DRH)
Compound 3.23-5 µM
Compound 3.23-2 µM
Compound 3.23-1 µM

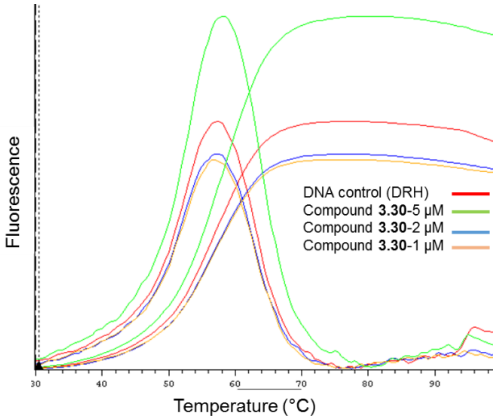

Fluorescence
Temperature (°C)
DNA control (DRH)
Compound 3.24-5 µM
Compound 3.24-2 µM
Compound 3.24-1 µM

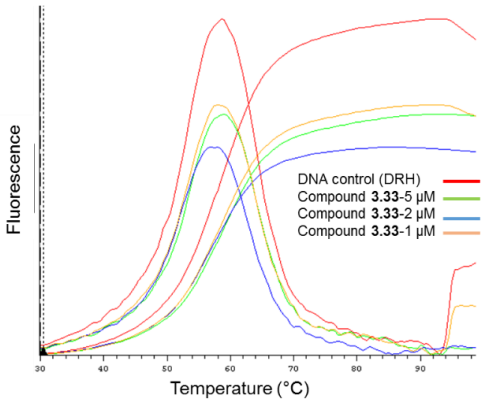

Fluorescence
Temperature (°C)
DNA control (DRH)
Compound 3.25-5 µM
Compound 3.25-2 µM
Compound 3.25-1 µM

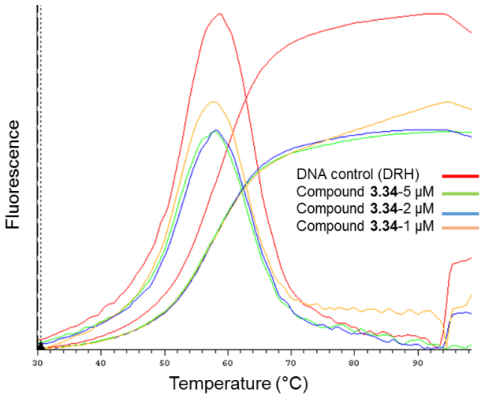

Fluorescence
DNA control (DRH)
Compound 3.27-5 µM
Compound 3.27-2 µM
Compound 3.27-1 µM
30
40
50
60
70
80
90
Temperature (°C)

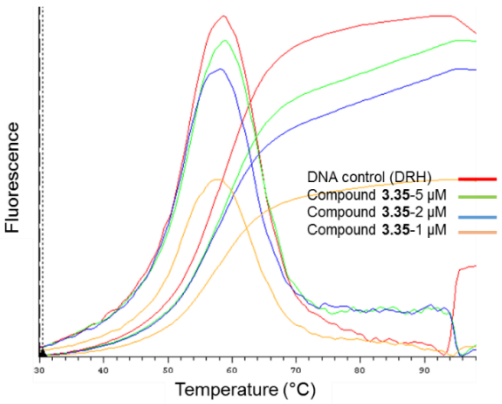

Fluorescence
DNA control (DRH)
Compound 3.28-5 µM
Compound 3.28-2 µM
Compound 3.28-1 µM
30
40
50
60
70
80
90
Temperature (°C)

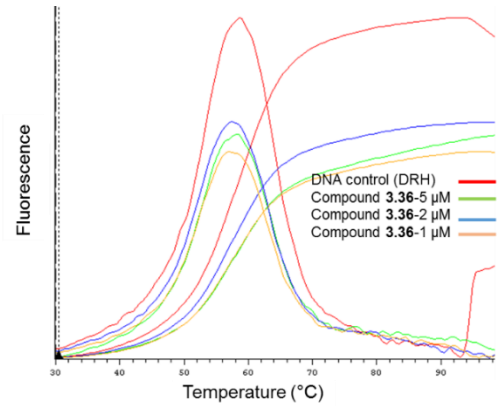

Fluorescence
DNA control (DRH)
Compound 3.29-5 µM
Compound 3.29-2 µM
Compound 3.29-1 µM
30
40
50
60
70
80
90
Temperature (°C)


Fluorescence
DNA control (DRH)
Compound 3.30-5 µM
Compound 3.30-2 µM
Compound 3.30-1 µM
30
40
50
60
70
80
90
Temperature (°C)

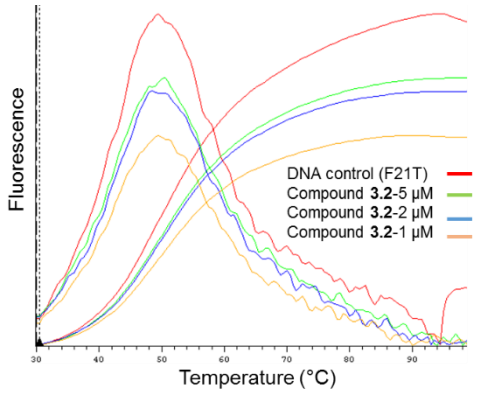

Fluorescence
DNA control (DRH)
Compound 3.33-5 µM
Compound 3.33-2 µM
Compound 3.33-1 µM
30
40
50
60
70
80
90
Temperature (°C)

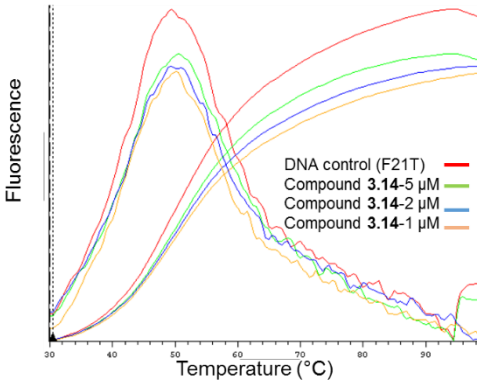

Fluorescence
DNA control (DRH)
Compound 3.34-5 µM
Compound 3.34-2 µM
Compound 3.34-1 µM
30
40
50
60
70
80
90
Temperature (°C)

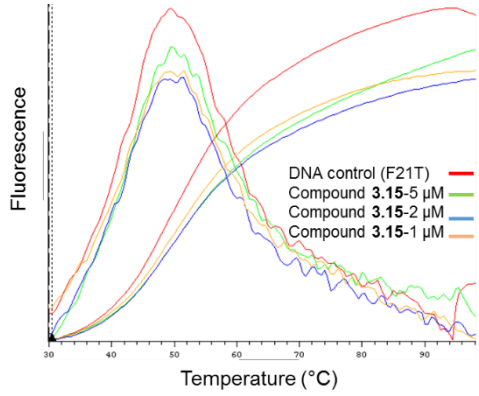

Fluorescence
DNA control (DRH)
Compound 3.35-5 µM
Compound 3.35-2 µM
Compound 3.35-1 µM
30
40
50
60
70
80
90
Temperature (°C)

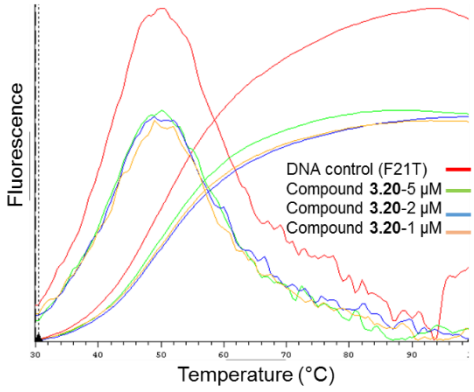

Fluorescence
DNA control (DRH)
Compound 3.36-5 µM
Compound 3.36-2 µM
Compound 3.36-1 µM
30
40
50
60
70
80
90
Temperature (°C)

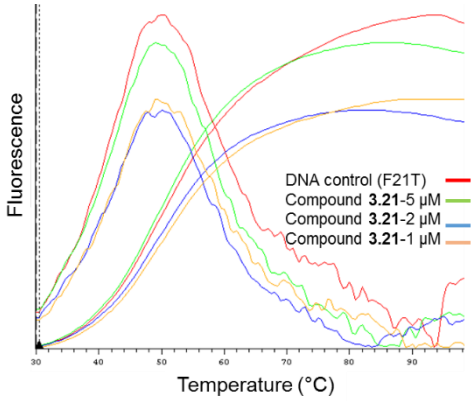
A2.1.2: Telomeric quadruplex (F21T) structure

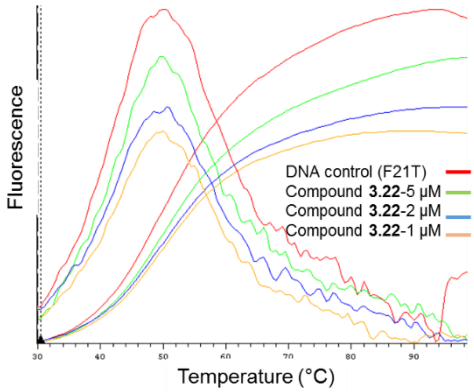


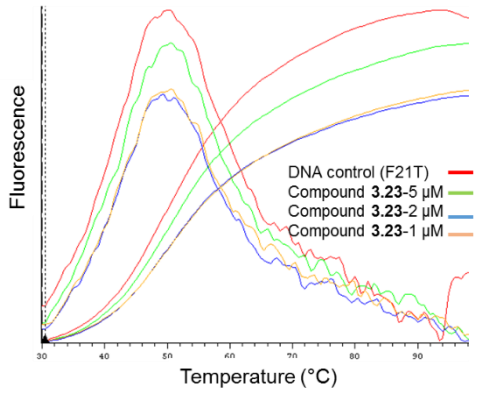


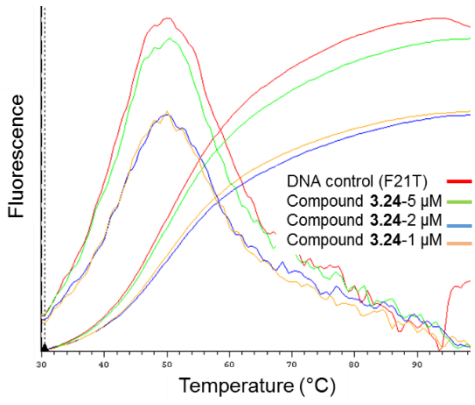


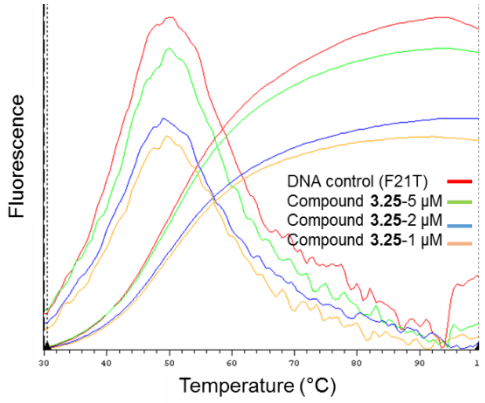


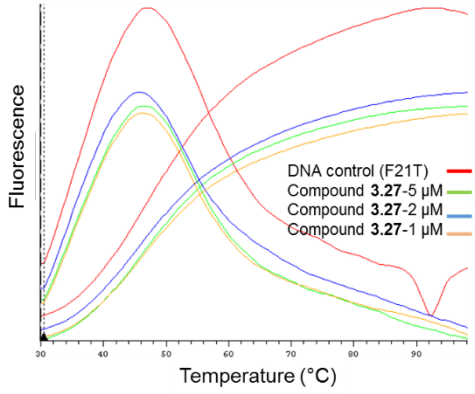


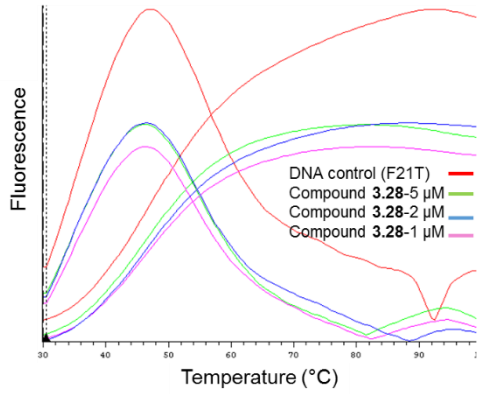


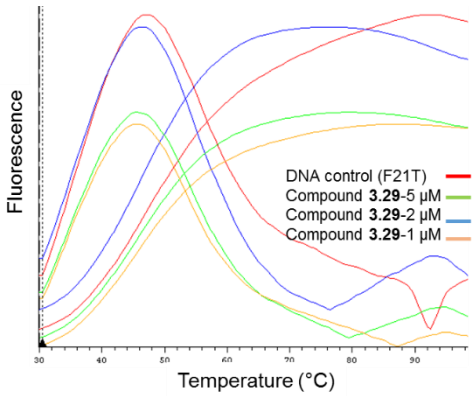


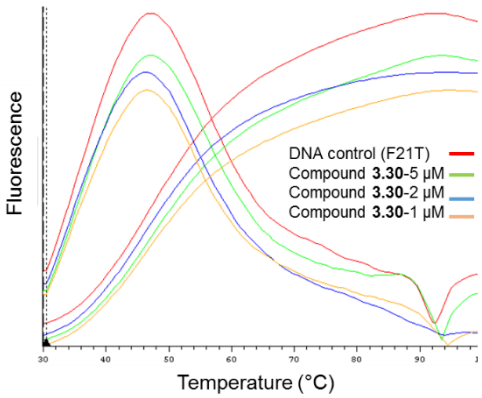


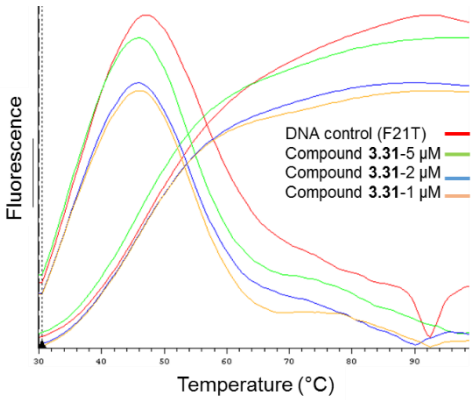

Fluorescence
DNA control (F21T)
Compound 3.25-5 μM
Compound 3.25-2 μM
Compound 3.25-1 μM
Temperature (°C)

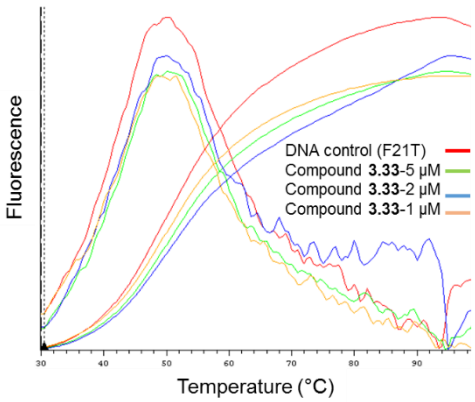

Fluorescence
DNA control (F21T)
Compound 3.27-5 μM
Compound 3.27-2 μM
Compound 3.27-1 μM
Temperature (°C)

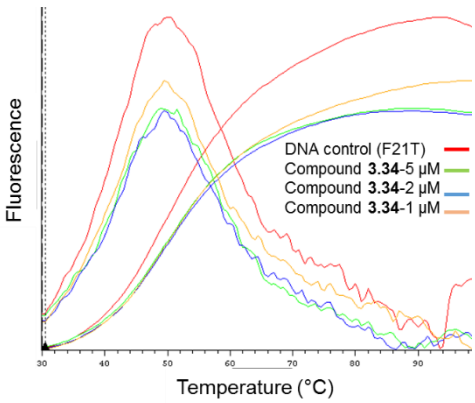

Fluorescence
DNA control (F21T)
Compound 3.28-5 μM
Compound 3.28-2 μM
Compound 3.28-1 μM
Temperature (°C)

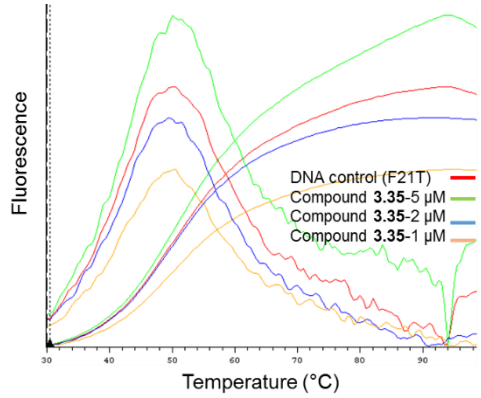

Fluorescence
DNA control (F21T)
Compound 3.29-5 μM
Compound 3.29-2 μM
Compound 3.29-1 μM
Temperature (°C)

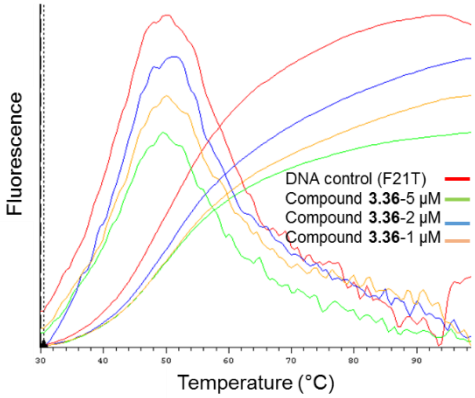

Fluorescence
DNA control (F21T)
Compound 3.30-5 μM
Compound 3.30-2 μM
Compound 3.30-1 μM
Temperature (°C)

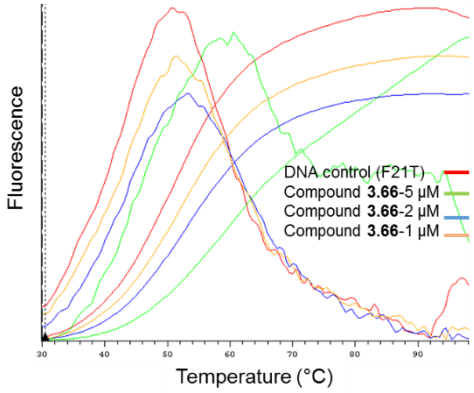

Fluorescence
DNA control (F21T)
Compound 3.31-5 μM
Compound 3.31-2 μM
Compound 3.31-1 μM
Temperature (°C)

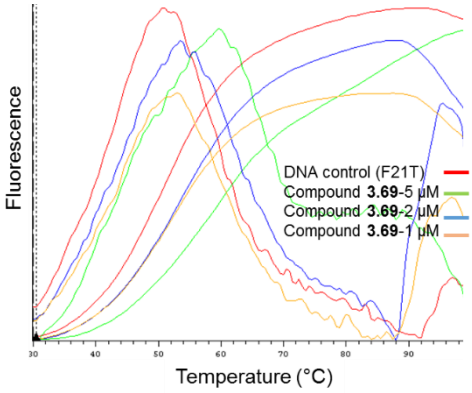

Fluorescence
DNA control (F21T)
Compound 3.33-5 μM
Compound 3.33-2 μM
Compound 3.33-1 μM
Temperature (°C)

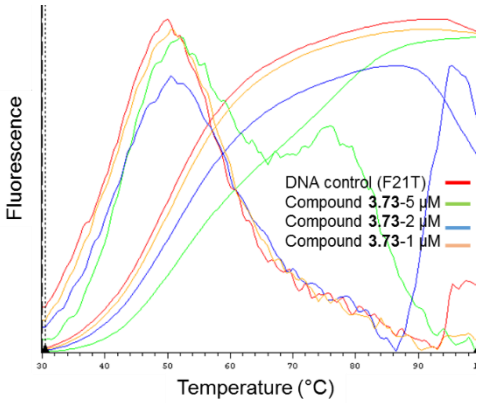

Fluorescence
DNA control (F21T)
Compound 3.34-5 μM
Compound 3.34-2 μM
Compound 3.34-1 μM
Temperature (°C)







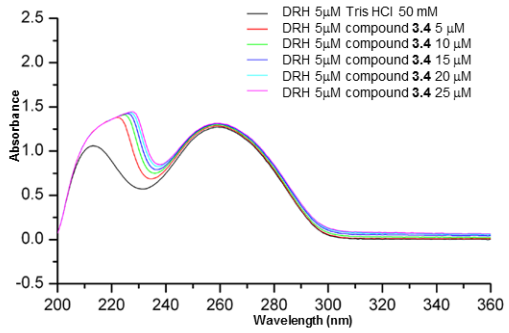


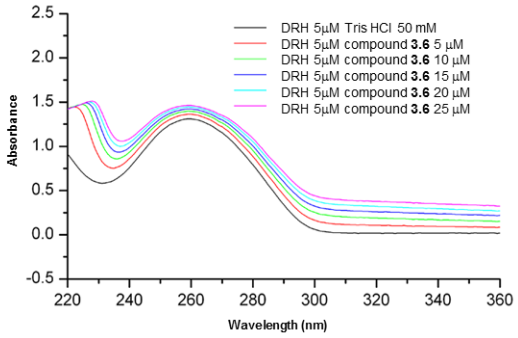


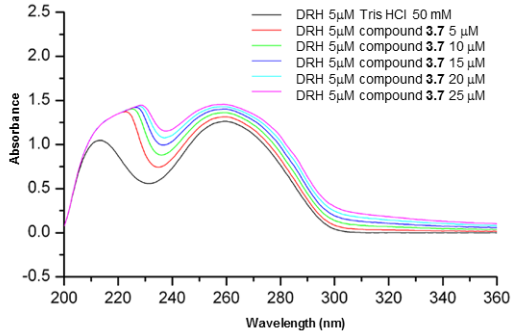
A2.2: UV calibration curves obtained on Circular Dichroism (CD) Studies

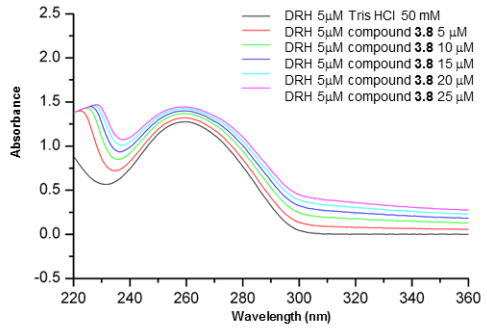
All CD spectrum of tested compounds were carried out for DNA/RNA hybrid duplex and telomeric quadruplex sequences. 50 mM Tris-HCl buffer (pH 7.4) was used to carry out the analysis with DRH oligonucleotide and 100 mM NaCl was used along with the Tris-HCl buffer for F21T sequence. Ligands were used at 0-5 equivalents over the oligonucleotides. In the case of CD analysis with DRH sequence 10 mm path length was applied, whereas 5 mm path length was applied for telomeric quadruplex sequence.

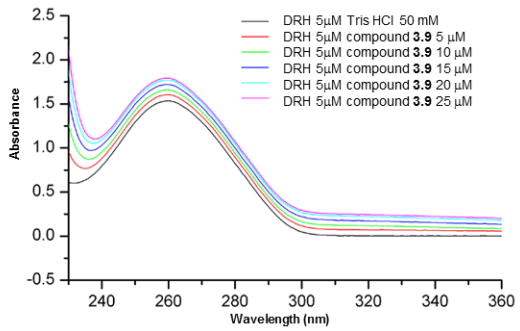
A2.2.1: DNA/RNA Hybrid duplex

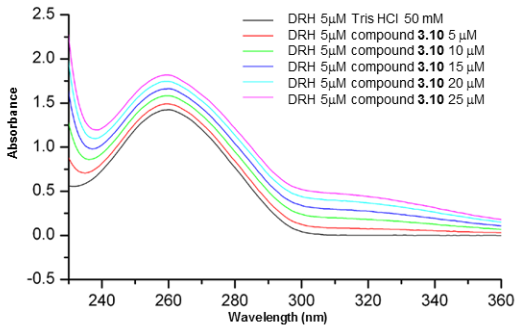


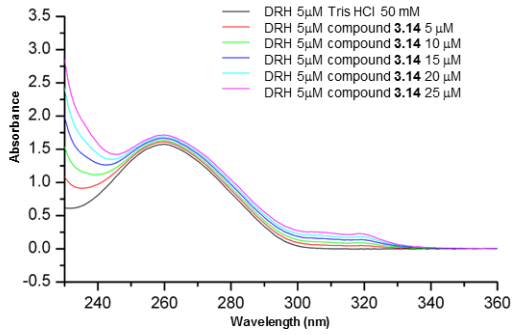


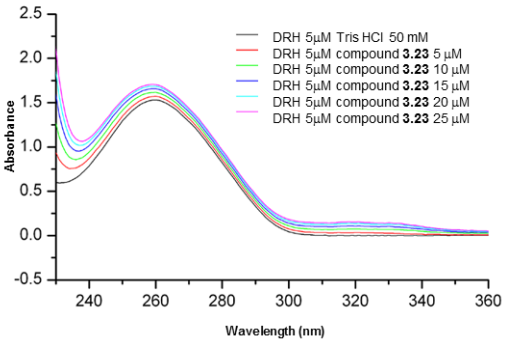


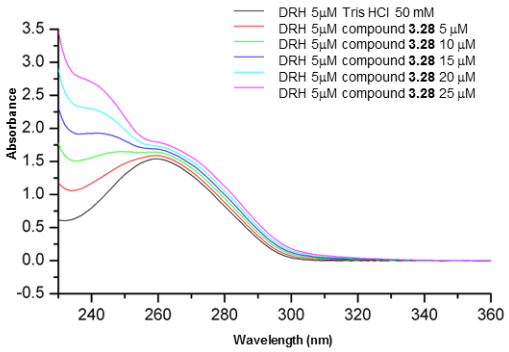


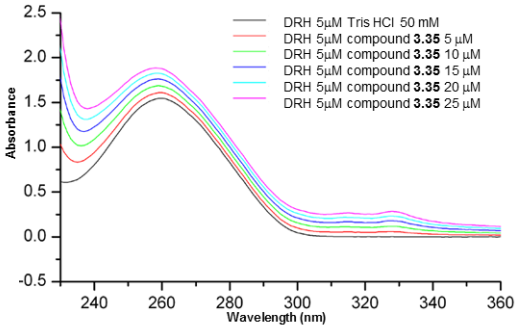




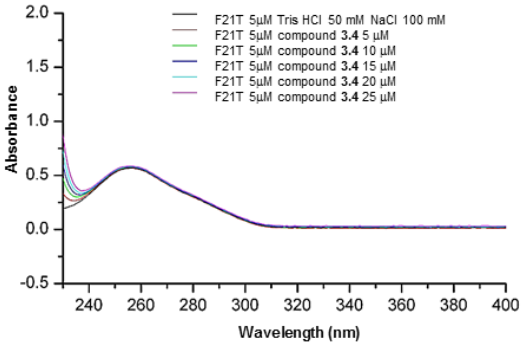


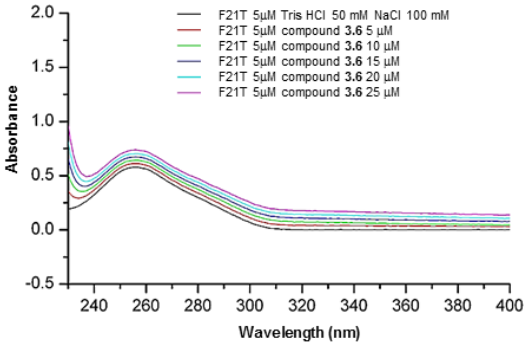


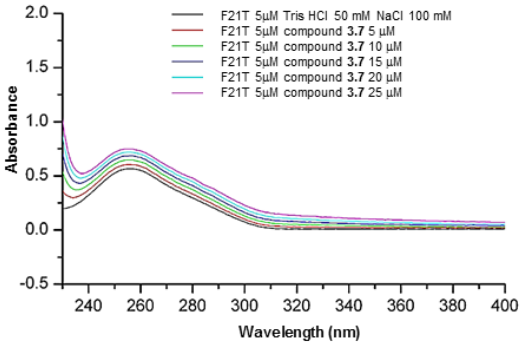


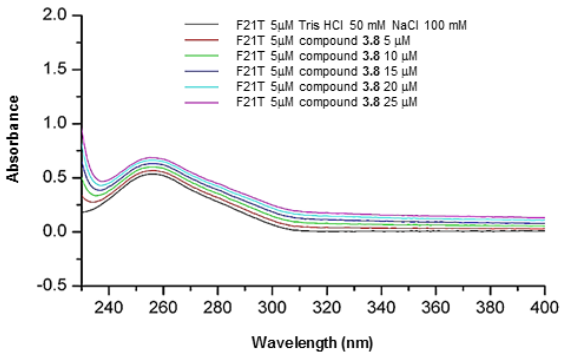


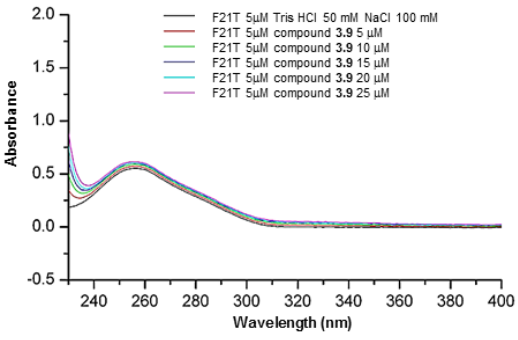
A2.2.2: Telomeric quadruplex (F21T) structure

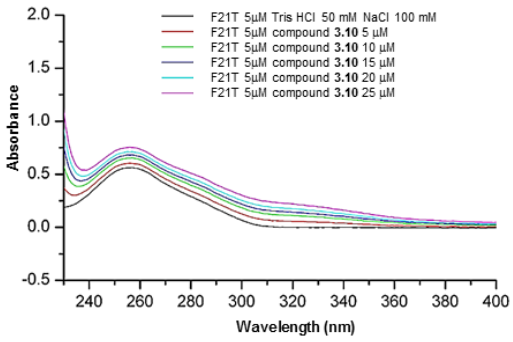


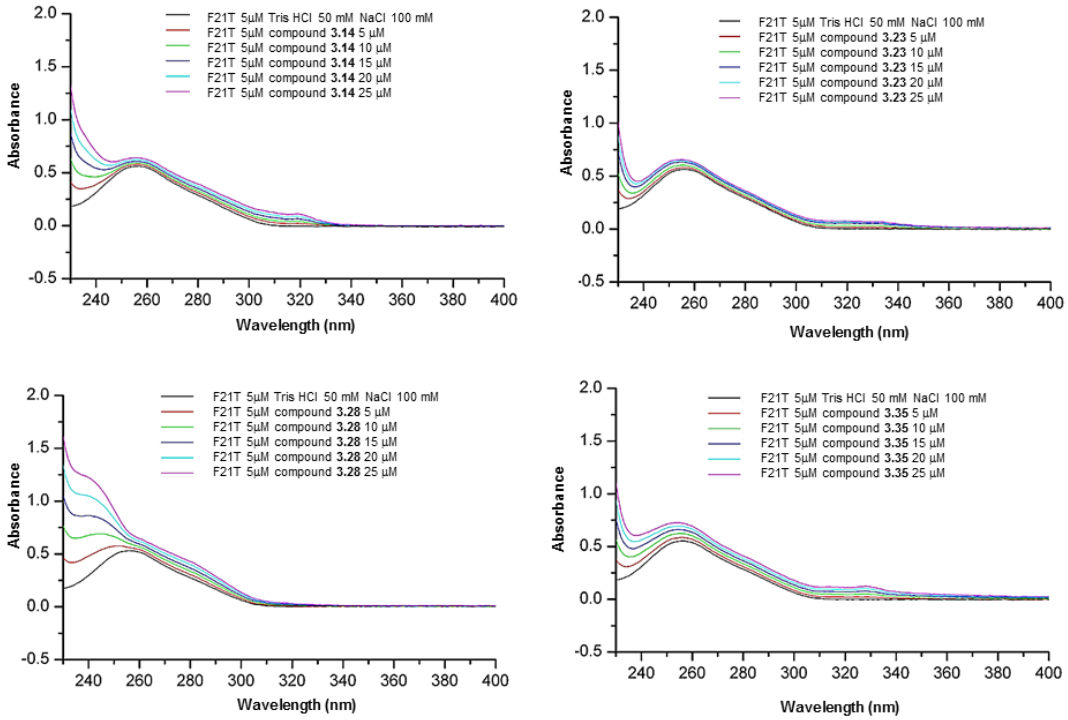






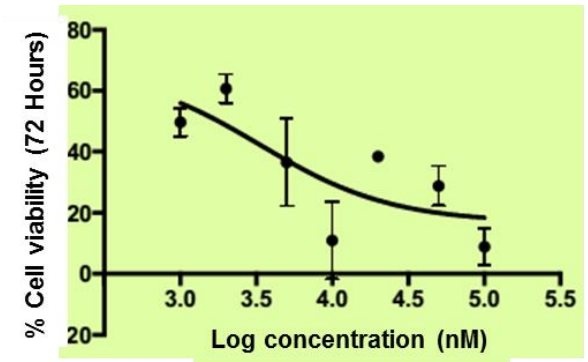




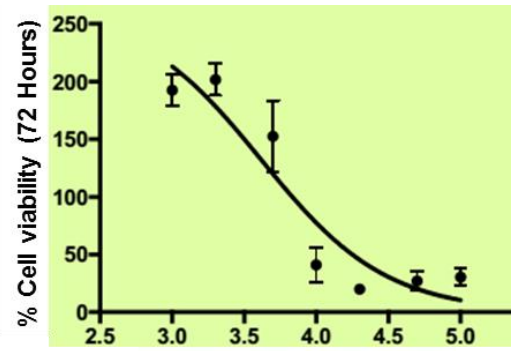


A2.3: IC_{50} curves observed in Cytotoxicity Studies (MTT Assay)

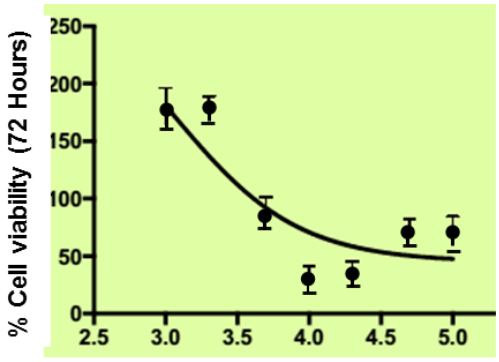
A2.3.1: MDA-MB-231 Cell Line (Breast Cancer Cell Line)



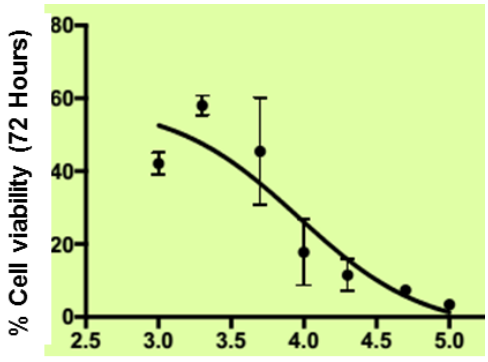
3.4 (IC_{50} = 3.4 μM, R^2 = 0.4007)



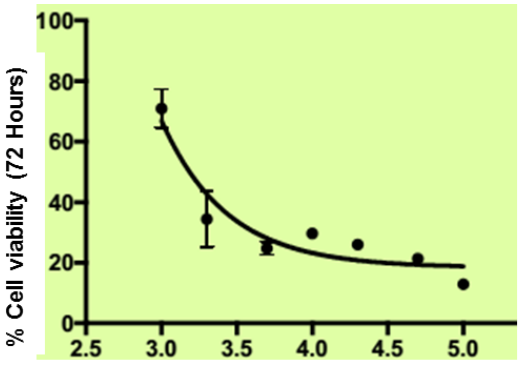
3.5 (IC_{50} = 4.2 μM, R^2 = 0.8256)



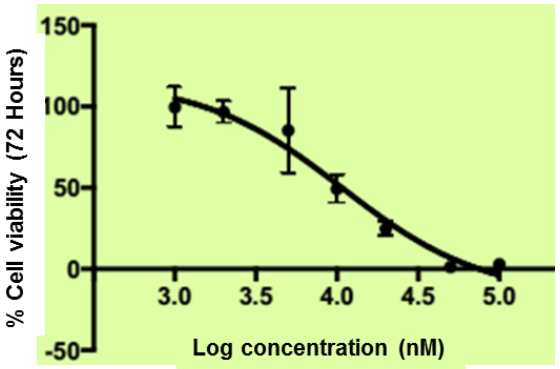
3.6 (IC_{50} = 1.1 μM, R^2 = 0.6440)



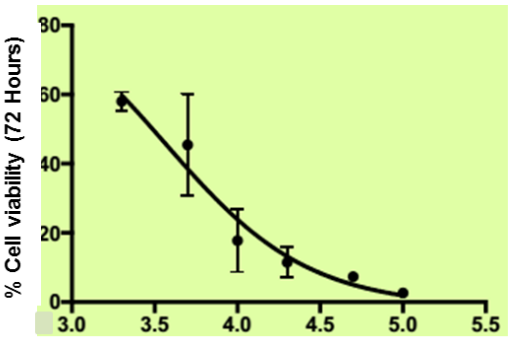
3.7 (IC_{50} = 9.2 μM, R^2 = 0.6933)



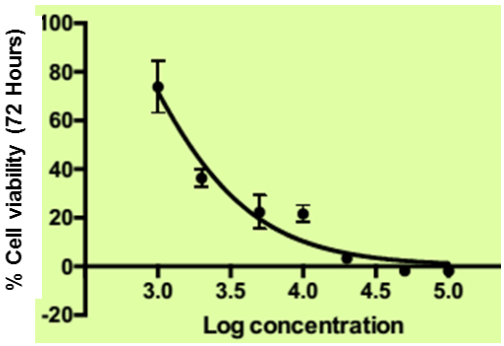
3.8 (IC_{50} = 4.2 μM, R^2 = 0.8018)



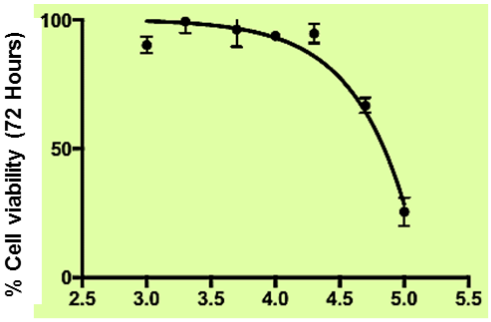
3.39 (IC_{50} = 10.8 μM, R^2 = 0.8266)



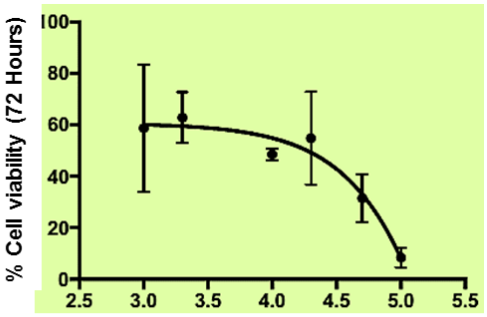
3.42 (IC_{50} = 3.6 μM, R^2 = 0.7530)



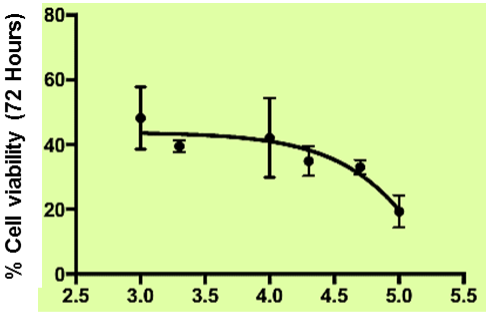
3.43 (IC_{50} = 0.5 μM, R^2 = 0.78758)



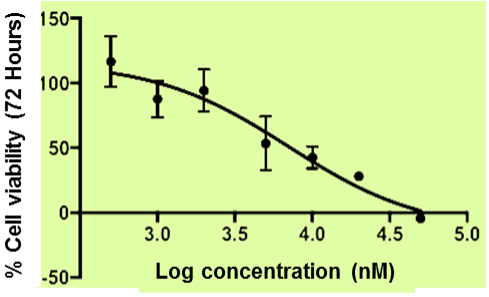
3.66 (IC_{50} = >100 μM, R^2 = 0.9029)



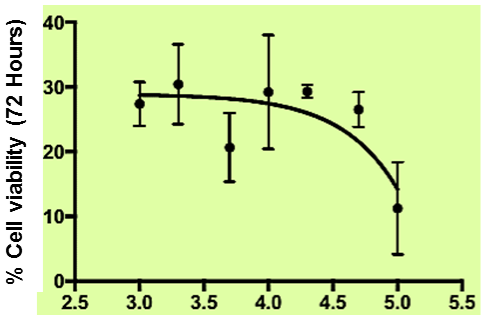
3.68 (IC_{50} = >100 μM, R^2 = 0.7067)



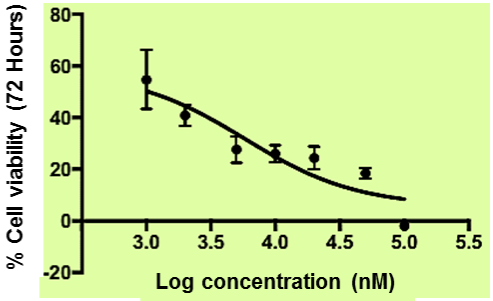
3.69 (IC_{50} = >100 μM, R^2 = 0.4862)



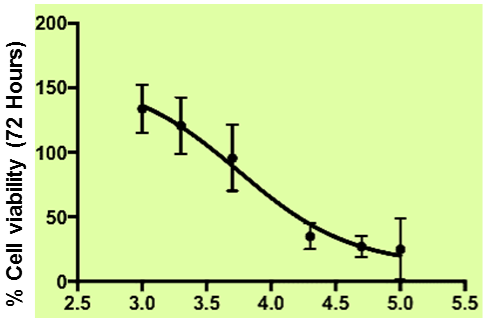
3.71 (IC_{50} = 15.1 μM, R^2 = 0.9163)



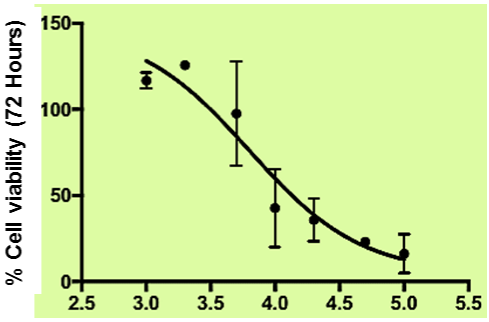
3.73 (IC_{50} = >100 μM, R^2 = 0.2428)



3.75 (IC_{50} = 5.8 μM, R^2 = 0.6942)

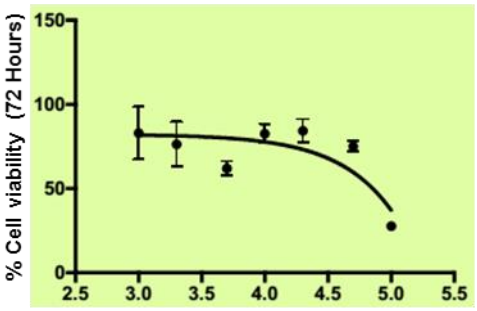


3.77 (IC_{50} = 5.6 μM, R^2 = 0.8815)

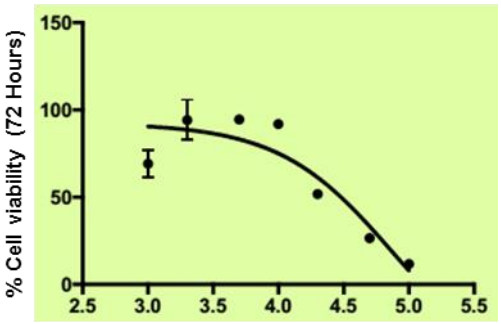


3.78 (IC_{50} = 6.3 μM, R^2 = 0.8615)

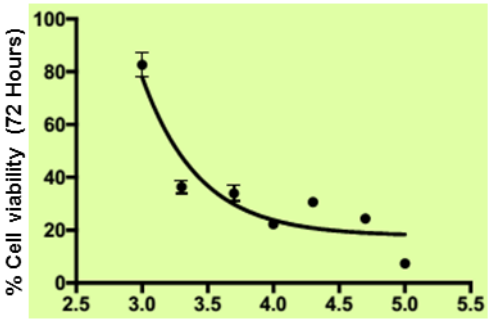
A2.3.2: HeLa Cell Line (Cervical Cancer Cell Line)



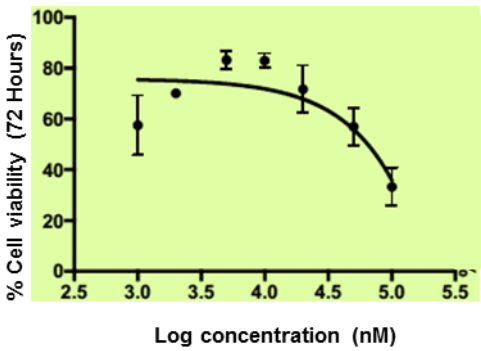
3.4 (IC_{50} = >100 μM, R^2 = 0.5789)



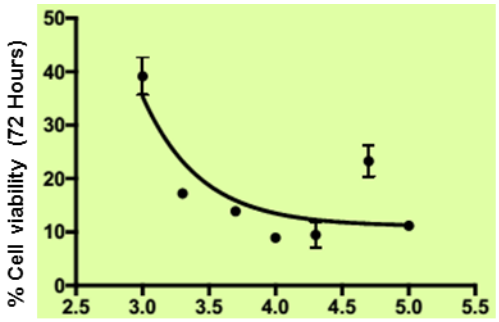
3.5 (IC_{50} = 75 μM, R^2 = 0.8320)



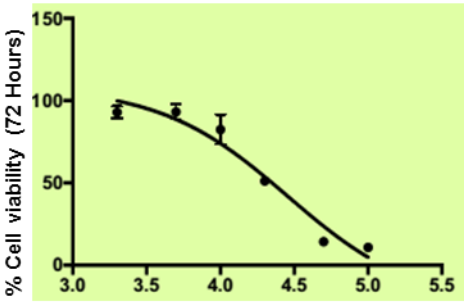
3.6 (IC_{50} = 4.2 μM, R^2 = 0.8496)



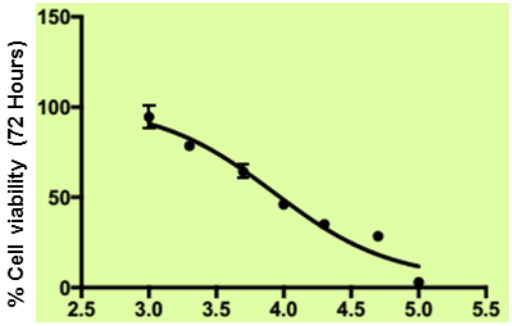
3.7 (IC_{50} = >100 μM, R^2 = 0.6079)



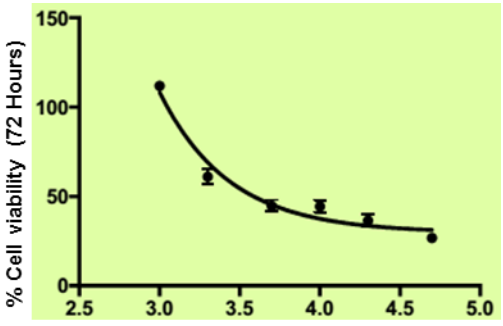
3.8 (IC_{50} = 2.7 µM, R^2 = 0.5789)



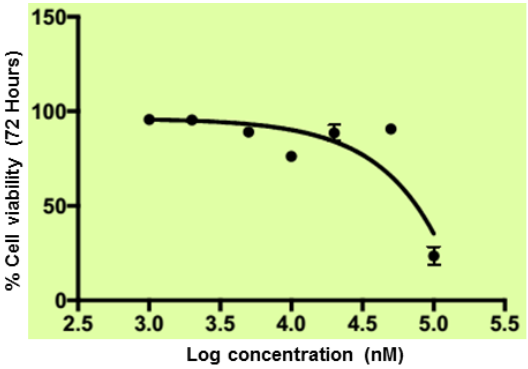
3.39 (IC_{50} = 28.6 µM, R^2 = 0.9311)



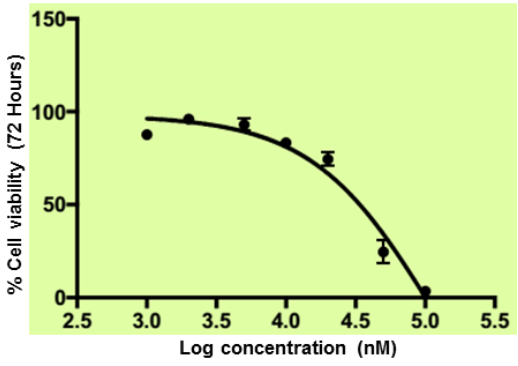
3.42 (IC_{50} = 8.5 µM, R^2 = 0.9413)



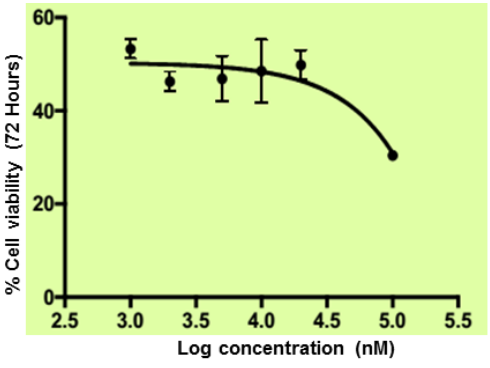
3.43 (IC_{50} = 1.2 µM, R^2 = 0.9429)



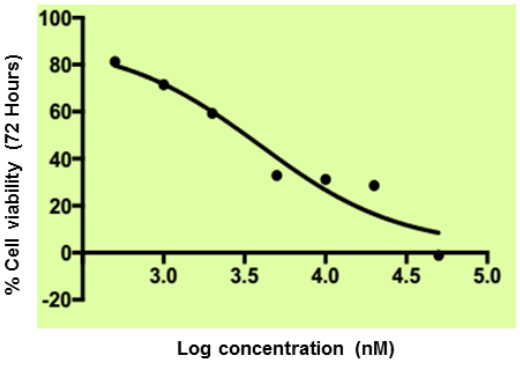
3.66 (IC_{50} = >100 µM, R^2 = 0.7295)



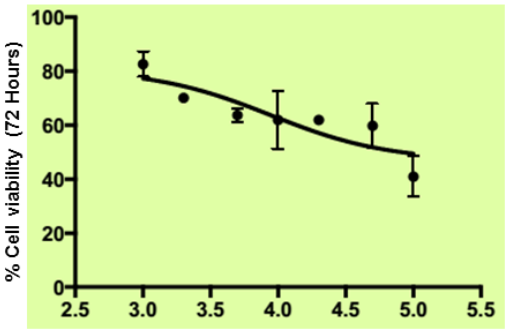
3.68 (IC_{50} = >100 µM, R^2 = 0.9524)



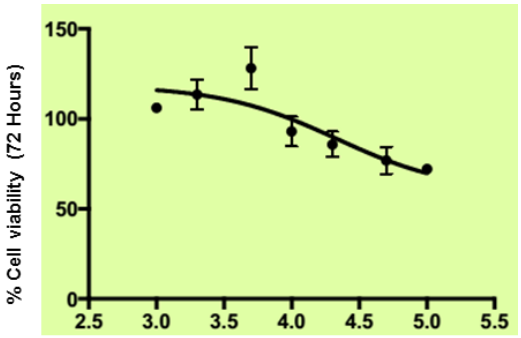
3.69 (IC_{50} = >100 µM, R^2 = 0.5601)



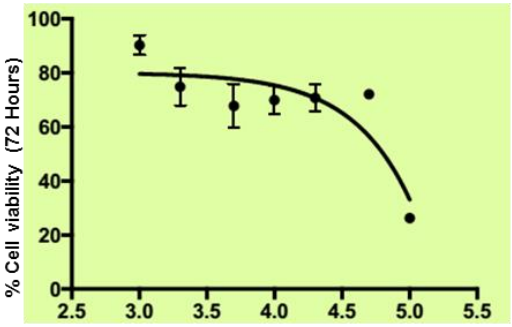
3.71 (IC_{50} = 3.9 µM, R^2 = 0.8554)



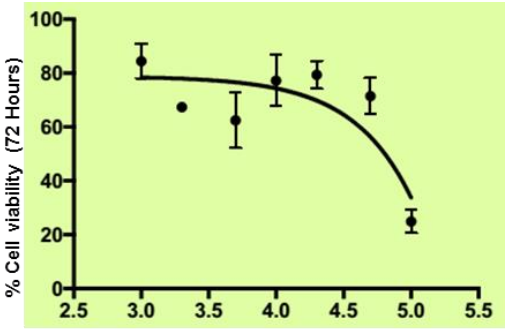
3.73 (IC_{50} = 8.8 µM, R^2 = 0.4796)



3.75 (IC_{50} = 21.5 µM, R^2 = 0.7190)

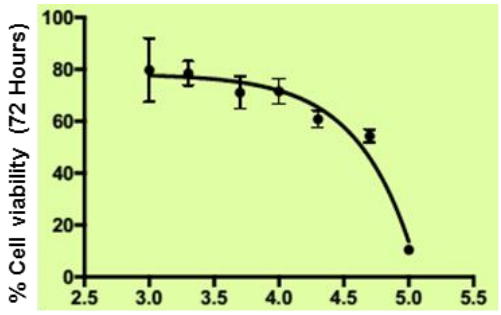


3.77 (IC_{50} = >100 µM, R^2 = 0.6579)

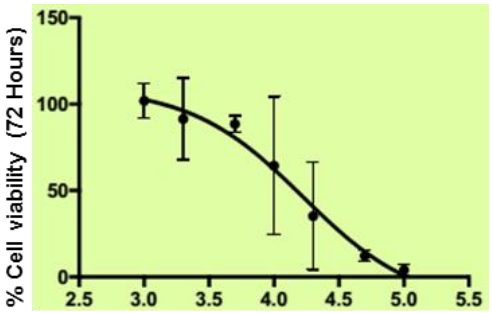


3.78 (IC_{50} = >100 µM, R^2 = 0.5343)

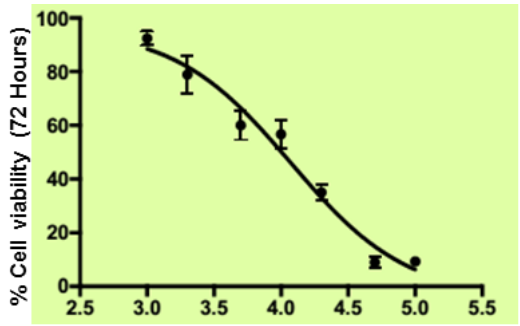
A2.3.2: NCI H1975 Cell Line (Non-Small Lung Cancer Cell Line)



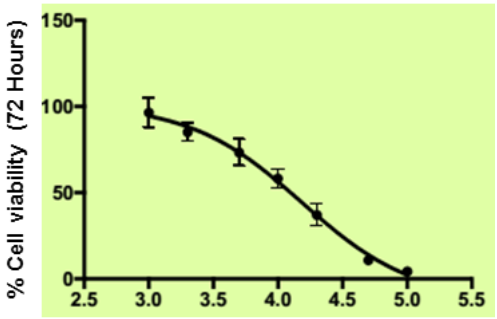
3.4 (IC_{50} = >100 µM, R^2 = 0.8420)



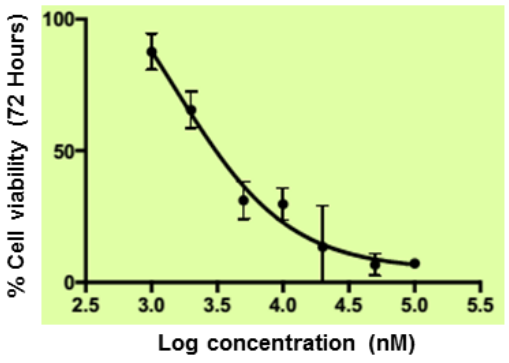
3.5 (IC_{50} = 16.8 µM, R^2 = 0.9623)



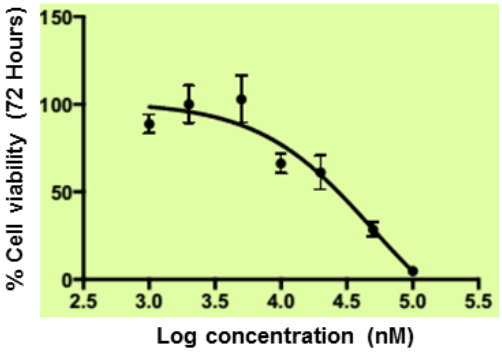
3.6 (IC_{50} = 11.8 µM, R^2 = 0.9623)



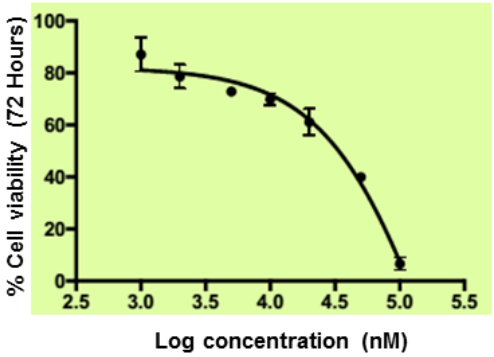
3.7 (IC_{50} = 15.3 µM, R^2 = 0.9398)



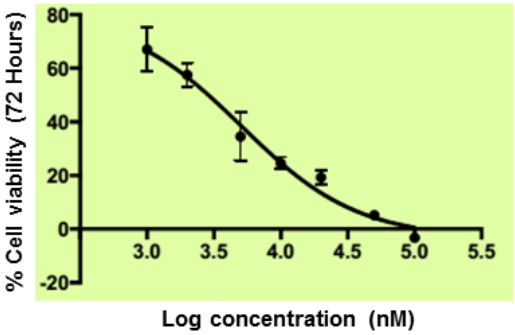
3.8 (IC_{50} = 1.5 µM, R^2 = 0.9343)



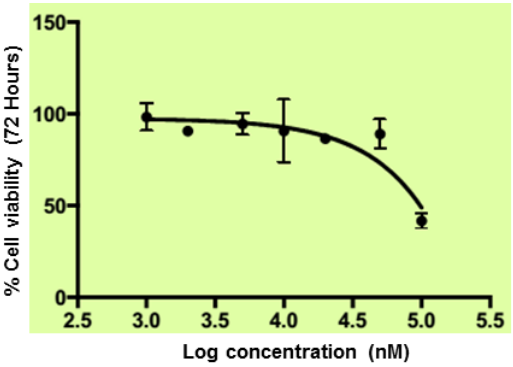
3.39 (IC_{50} = 49.3 µM, R^2 = 0.8490)



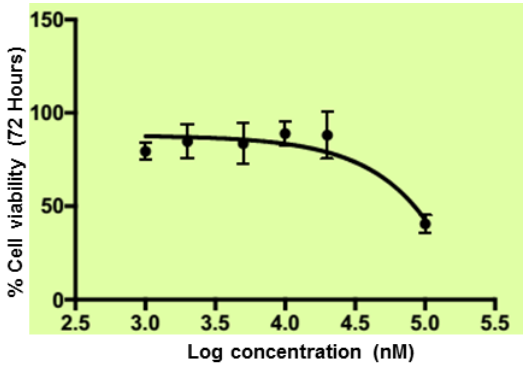
3.42 (IC_{50} = >100 µM, R^2 = 0.9425)



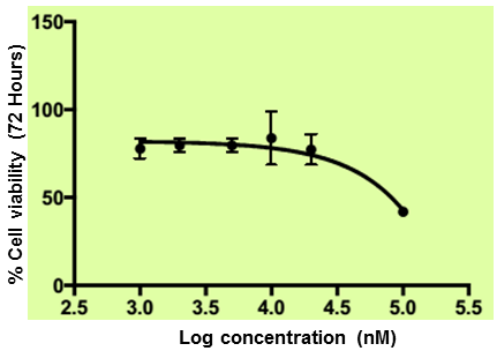
3.43 (IC_{50} = 5.1 µM, R^2 = 0.9006)



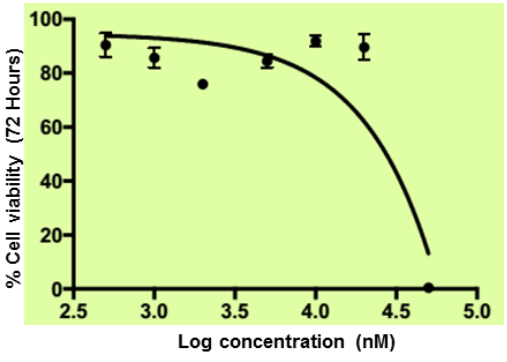
3.66 (IC_{50} = >100 µM, R^2 = 0.5931)



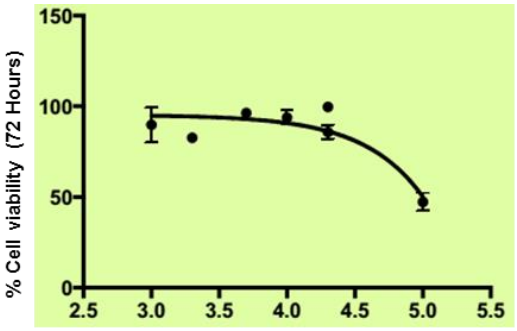
3.68 (IC_{50} = >100 µM, R^2 = 0.5776)



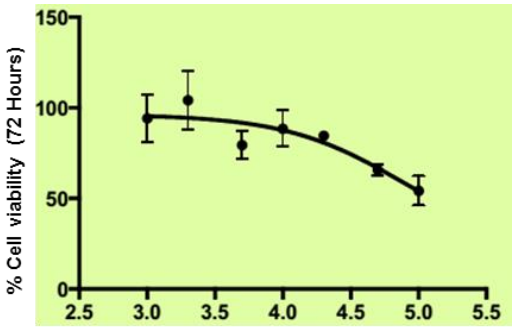
3.69 (IC_{50} = >100 µM, R^2 = 0.6982)



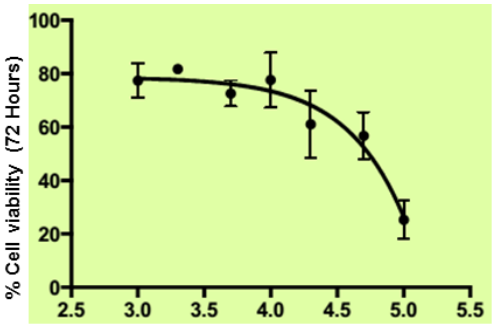
3.71 (IC_{50} = >100 µM, R^2 = 0.7663)



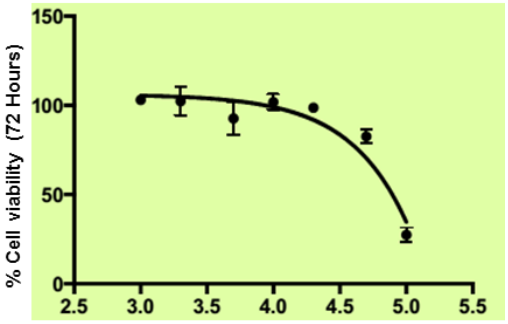
3.73 (IC_{50} = >100 μM, R^2 = 0.6796)



3.75 (IC_{50} = 75 μM, R^2 = 0.6732)



3.77 (IC_{50} = >100 μM, R^2 = 0.6795)



3.78 (IC_{50} = >100 μM, R^2 = 0.8393)

References

1. Hanahan, D., and Weinberg, R. A. (2011) Hallmarks of cancer: the next generation, *Cell 144*, 646-674.
2. Schroeder, A., Heller, D. A., Winslow, M. M., Dahlman, J. E., Pratt, G. W., Langer, R., Jacks, T., and Anderson, D. G. (2012) Treating metastatic cancer with nanotechnology, *Nat Rev Cancer 12*, 39-50.
3. 'Cancer Fact sheet N°297'. World Health Organization. February 2015. http://www.who.int/mediacentre/factsheets/fs297/en/ (Retrieved Nov 2015)).
4. Dollinger, M., Rosenbaum,E., and Cable, G. (1997) Everyone's Guide to Cancer Therapy (Kansas City: Andrews McMeel Publishing).
5. Hartwell, L. (1992) Defects in a Cell-Cycle Checkpoint May Be Responsible for the Genomic Instability of Cancer-Cells, *Cell 71*, 543-546.
6. Loeb, L. A. (1991) Mutator Phenotype May Be Required For Multistage Carcinogenesis, *Cancer Research 51*, 3075-3079.
7. Vorlickova, M., Chladkova, J., Kejnovska, I., Fialova, M., and Kypr, J. (2005) Guanine tetraplex topology of human telomere DNA is governed by the number of (TTAGGG) repeats, *Nucleic Acids Res 33*, 5851-5860.
8. Hanahan, D., and Weinberg, R. A. (2000) The hallmarks of cancer, *Cell 100*, 57-70.
9. Jiang, B. H., and Liu, L. Z. (2009) PI3K/PTEN signaling in angiogenesis and tumorigenesis, *Advances in cancer research 102*, 19-65.
10. Ikushima, H., and Miyazono, K. (2010) TGFbeta signalling: a complex web in cancer progression, *Nat Rev Cancer 10*, 415-424.
11. Davies, M. A., and Samuels, Y. (2010) Analysis of the genome to personalize therapy for melanoma, *Oncogene 29*, 5545-5555.
12. DeVita, V. T., Jr., and Chu, E. (2008) A history of cancer chemotherapy, *Cancer Res 68*, 8643-8653.
13. Yap, T. A., Omlin, A., and de Bono, J. S. (2013) Development of therapeutic combinations targeting major cancer signaling pathways, *Journal of clinical oncology : official journal of the American Society of Clinical Oncology 31*, 1592-1605.
14. Hurley, L. H. (2002) DNA and its associated processes as targets for cancer therapy, *Nat Rev Cancer 2*, 188-200.
15. Lerman, L. S. (1961) Structural considerations in the interaction of DNA and acridines, *Journal of molecular biology 3*, 18-30.
16. Chaires, J. B. (1997) Energetics of drug-DNA interactions, *Biopolymers 44*, 201-215.
17. Drygin, D., Siddiqui-Jain, A., O'Brien, S., Schwaebe, M., Lin, A., Bliesath, J., Ho, C. B., Proffitt, C., Trent, K., Whitten, J. P., Lim, J. K., Von Hoff, D., Anderes, K., and Rice, W. G. (2009) Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis, *Cancer Res 69*, 7653-7661.

18. Balasubramanian, S., Hurley, L. H., and Neidle, S. (2011) Targeting G-quadruplexes in gene promoters: a novel anticancer strategy, *Nature reviews. Drug discovery 10*, 261-275.
19. Nussbaumer, S., Bonnabry, P., Veuthey, J. L., and Fleury-Souverain, S. (2011) Analysis of anticancer drugs: A review, *Talanta 85*, 2265-2289.
20. Peters, G. J. (2014) Novel developments in the use of antimetabolites, *Nucleosides, nucleotides & nucleic acids 33*, 358-374.
21. Calderon-Montano, J. M., Burgos-Moron, E., Orta, M. L., and Lopez-Lazaro, M. (2014) Effect of DNA repair deficiencies on the cytotoxicity of drugs used in cancer therapy - a review, *Current medicinal chemistry 21*, 3419-3454.
22. Jordan, M. A. (2002) Mechanism of action of antitumor drugs that interact with microtubules and tubulin, *Current medicinal chemistry. Anti-cancer agents 2*, 1-17.
23. Ganguly, A., Yang, H., and Cabral, F. (2010) Paclitaxel-dependent cell lines reveal a novel drug activity, *Molecular cancer therapeutics 9*, 2914-2923.
24. Okouneva, T., Hill, B. T., Wilson, L., and Jordan, M. A. (2003) The effects of vinflunine, vinorelbine, and vinblastine on centromere dynamics, *Molecular cancer therapeutics 2*, 427-436.
25. Huang, M., Shen, A., Ding, J., and Geng, M. (2014) Molecularly targeted cancer therapy: some lessons from the past decade, *Trends in Pharmacological Sciences 35*, 41-50.
26. Janne, P. A., Gray, N., and Settleman, J. (2009) Factors underlying sensitivity of cancers to small-molecule kinase inhibitors, *Nature reviews. Drug discovery 8*, 709-723.
27. Downward, J. (2003) Targeting RAS signalling pathways in cancer therapy, *Nat Rev Cancer 3*, 11-22.
28. Dickson, M. A., and Schwartz, G. K. (2009) Development of cell-cycle inhibitors for cancer therapy, *Current Oncology 16*, 36-43.
29. Bonvini, P., Zorzi, E., Basso, G., and Rosolen, A. (2007) Bortezomib-mediated 26S proteasome inhibition causes cell-cycle arrest and induces apoptosis in CD-30+ anaplastic large cell lymphoma, *Leukemia 21*, 838-842.
30. Jones, A., Barrett-Lee, P. J., and Mason, M. D. (2004) Hormonal therapy in cancer, *Medicine 32*, 30-32.
31. Early Breast Cancer Trialists' Collaborative, G. (1998) Tamoxifen for early breast cancer: an overview of the randomised trials, *The Lancet 351*, 1451-1467.
32. Abraham, J., and Staffurth, J. (2016) Hormonal therapy for cancer, *Medicine 44*, 30-33.
33. Thomas, A., Teicher, B. A., and Hassan, R. (2016) Antibody-drug conjugates for cancer therapy, *The Lancet. Oncology 17*, e254-262.
34. Tan, A. R., and Swain, S. M. (2003) Ongoing adjuvant trials with trastuzumab in breast cancer, *Seminars in oncology 30*, 54-64.
35. Kelland, L. R. (2005) Overcoming the immortality of tumour cells by telomere and telomerase based cancer therapeutics--current status and future prospects, *Eur J Cancer 41*, 971-979.

36. Watson, J. D., and Crick, F. H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid, *Nature 171*, 737-738.
37. Rich, A. (1993) DNA comes in many forms, *Gene 135*, 99-109.
38. Mills, M., Lacroix, L., Arimondo, P. B., Leroy, J. L., Francois, J. C., Klump, H., and Mergny, J. L. (2002) Unusual DNA conformations: implications for telomeres, *Current medicinal chemistry. Anti-cancer agents 2*, 627-644.
39. Ho, P. S., and Carter, M. (2011) *DNA Replication-Current Advances, Chapter 1- DNA Structure: Alphabet Soup for the Cellular Soul*.
40. Mandelkern, M., Elias, J. G., Eden, D., and Crothers, D. M. (1981) The dimensions of DNA in solution, *Journal of molecular biology 152*, 153-161.
41. Yakovchuk, P., Protozanova, E., and Frank-Kamenetskii, M. D. (2006) Base-stacking and base-pairing contributions into thermal stability of the DNA double helix, *Nucleic Acids Research 34*, 564-574.
42. Alberts B, J. A., Lewis J, . (2002) The Structure and Function of DNA, *Molecular Biology of the Cell. 4th edition. New York: Garland Science*.
43. Hays, F. A., Teegarden, A., Jones, Z. J., Harms, M., Raup, D., Watson, J., Cavaliere, E., and Ho, P. S. (2005) How sequence defines structure: a crystallographic map of DNA structure and conformation, *Proceedings of the National Academy of Sciences of the United States of America 102*, 7157-7162.
44. Wang, A. H. J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., Vanboom, J. H., Vandermarel, G., and Rich, A. (1979) Molecular-Structure of a Left-Handed Double Helical DNA Fragment at Atomic Resolution, *Nature 282*, 680-686.
45. Prive, G. G., Yanagi, K., and Dickerson, R. E. (1991) Structure of the B-DNA decamer C-C-A-A-C-G-T-T-G-G and comparison with isomorphous decamers C-C-A-A-G-A-T-T-G-G and C-C-A-G-G-C-C-T-G-G, *Journal of molecular biology 217*, 177-199.
46. Richmond, T. J., and Davey, C. A. (2003) The structure of DNA in the nucleosome core, *Nature 423*, 145-150.
47. Haschemeyer, A. E. V., and Rich, A. (1967) Nucleoside conformations: an analysis of steric barriers to rotation about the glycosidic bond, *Journal of molecular biology 27*, 369-384.
48. Arnott, S., and Selsing, E. (1974) Structures for the polynucleotide complexes poly(dA) with poly (dT) and poly(dT) with poly(dA) with poly (dT), *Journal of molecular biology 88*, 509-521.
49. Morgan, A. R., and Wells, R. D. (1968) Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences, *Journal of molecular biology 37*, 63-80.
50. Chamberl.Mj, and Patterso.Dl. (1965) Physical and Chemical Characterization of Ordered Complexes Formed between Polyinosinic Acid Polycytidylic Acid and Their Deoxyribo-Analogues, *Journal of molecular biology 12*, 410-&.

51. Felsenfeld, G., Davies, D. R., and Rich, A. (1957) Formation of a 3-Stranded Polynucleotide Molecule, *Journal of the American Chemical Society 79*, 2023-2024.
52. Vasquez, K. M., and Wilson, J. H. (1998) Triplex-directed modification of genes and gene activity, *Trends Biochem Sci 23*, 4-9.
53. Raghavan, S. C., Chastain, P., Lee, J. S., Hegde, B. G., Houston, S., Langen, R., Hsieh, C. L., Haworth, I. S., and Lieber, M. R. (2005) Evidence for a triplex DNA conformation at the bcl-2 major breakpoint region of the t(14;18) translocation, *J Biol Chem 280*, 22749-22760.
54. Dayn, A., Samadashwily, G. M., and Mirkin, S. M. (1992) Intramolecular DNA triplexes: unusual sequence requirements and influence on DNA polymerization, *Proceedings of the National Academy of Sciences of the United States of America 89*, 11406-11410.
55. Knauert, M. P., and Glazer, P. M. (2001) Triplex forming oligonucleotides: sequence-specific tools for gene targeting, *Human Molecular Genetics 10*, 2243-2251.
56. Rich, A. (1958) The molecular structure of polyinosinic acid, *Biochim Biophys Acta 29*, 502-509.
57. Zimmerman, S. B., Cohen, G. H., and Davies, D. R. (1975) X-ray fiber diffraction and model-building study of polyguanylic acid and polyinosinic acid, *Journal of molecular biology 92*, 181-192.
58. Mirkin, S. M. (2008) Discovery of alternative DNA structures: a heroic decade (1979-1989), *Front Biosci 13*, 1064-1071.
59. Tera, M., Iida, K., Ishizuka, H., Takagi, M., Suganuma, M., Doi, T., Shin-ya, K., and Nagasawa, K. (2009) Synthesis of a potent G-quadruplex-binding macrocyclic heptaoxazole, *Chembiochem 10*, 431-435.
60. Williamson, J. R., Raghuraman, M. K., and Cech, T. R. (1989) Monovalent cation-induced structure of telomeric DNA: the G-quartet model, *Cell 59*, 871-880.
61. Bochman, M. L., Paeschke, K., and Zakian, V. A. (2012) DNA secondary structures: stability and function of G-quadruplex structures, *Nature reviews. Genetics 13*, 770-780.
62. Murat, P., Singh, Y., and Defrancq, E. (2011) Methods for investigating G-quadruplex DNA/ligand interactions, *Chem Soc Rev 40*, 5293-5307.
63. Gehring, K., Leroy, J. L., and Gueron, M. (1993) A tetrameric DNA structure with protonated cytosine.cytosine base pairs, *Nature 363*, 561-565.
64. Gueron, M., and Leroy, J. L. (2000) The i-motif in nucleic acids, *Curr Opin Struct Biol 10*, 326-331.
65. Choi, J., Kim, S., Tachikawa, T., Fujitsuka, M., and Majima, T. (2011) pH-induced intramolecular folding dynamics of i-motif DNA, *J Am Chem Soc 133*, 16146-16153.
66. Huppert, J. L., and Balasubramanian, S. (2007) G-quadruplexes in promoters throughout the human genome, *Nucleic Acids Res 35*, 406-413.

67. Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K., and Dickerson, R. E. (1980) Crystal structure analysis of a complete turn of B-DNA, *Nature 287*, 755-758.
68. Burley, S. K. (1996) X-ray crystallographic studies of eukaryotic transcription initiation factors, *Philosophical transactions of the Royal Society of London. Series B, Biological sciences 351*, 483-489.
69. Schluth-Bolard, C., Ottaviani, A., Bah, A., Boussouar, A., Gilson, E., and Magdinier, F. (2011) Dynamics and plasticity of chromosome ends: consequences in human pathologies, *Atlas of Genetics and Cytogenetics in Oncology and Haematology 14*, 501-524.
70. de Lange, T., Shiue, L., Myers, R. M., Cox, D. R., Naylor, S. L., Killery, A. M., and Varmus, H. E. (1990) Structure and variability of human chromosome ends, *Molecular and cellular biology 10*, 518-527.
71. Harley, C. B., Futcher, A. B., and Greider, C. W. (1990) Telomeres shorten during ageing of human fibroblasts, *Nature 345*, 458-460.
72. Lindsey, J., McGill, N. I., Lindsey, L. A., Green, D. K., and Cooke, H. J. (1991) In vivo loss of telomeric repeats with age in humans, *Mutation research 256*, 45-48.
73. Hayflick, L. (1965) The Limited in Vitro Lifetime of Human Diploid Cell Strains, *Experimental cell research 37*, 614-636.
74. Shay, J. W., Pereirasmith, O. M., and Wright, W. E. (1991) A Role for Both Rb and P53 in the Regulation of Human Cellular Senescence, *Experimental cell research 196*, 33-39.
75. Wright, W. E., Pereira-Smith, O. M., and Shay, J. W. (1989) Reversible cellular senescence: implications for immortalization of normal human diploid fibroblasts, *Molecular and cellular biology 9*, 3088-3092.
76. Shay, J. W., Wright, W. E., Brasiskyte, D., and Van der Haegen, B. A. (1993) E6 of human papillomavirus type 16 can overcome the M1 stage of immortalization in human mammary epithelial cells but not in human fibroblasts, *Oncogene 8*, 1407-1413.
77. Shay, J. W., and Wright, W. E. (1989) Quantitation of the frequency of immortalization of normal human diploid fibroblasts by SV40 large T-antigen, *Experimental cell research 184*, 109-118.
78. Counter, C. M., Avilion, A. A., Lefeuvre, C. E., Stewart, N. G., Greider, C. W., Harley, C. B., and Bacchetti, S. (1992) Telomere Shortening Associated with Chromosome Instability Is Arrested in Immortal Cells Which Express Telomerase Activity, *Embo J 11*, 1921-1929.
79. Greider, C. W., and Blackburn, E. H. (1985) Identification of a specific telomere terminal transferase activity in Tetrahymena extracts, *Cell 43*, 405-413.
80. Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L., Coviello, G. M., Wright, W. E., Weinrich, S. L., and Shay, J. W. (1994) Specific association of human telomerase activity with immortal cells and cancer, *Science (New York, N.Y.) 266*, 2011-2015.

81. Shay, J. W., and Bacchetti, S. (1997) A survey of telomerase activity in human cancer, *Eur J Cancer 33*, 787-791.
82. Moyzis, R. K., Buckingham, J. M., Cram, L. S., Dani, M., Deaven, L. L., Jones, M. D., Meyne, J., Ratliff, R. L., and Wu, J. R. (1988) A Highly Conserved Repetitive DNA-Sequence, (Ttaggg)N, Present at the Telomeres of Human-Chromosomes, *Proceedings of the National Academy of Sciences of the United States of America 85*, 6622-6626.
83. Reddel, R. R. (2014) Telomere maintenance mechanisms in cancer: clinical implications, *Curr Pharm Des 20*, 6361-6374.
84. Griffith, J. D., Comeau, L., Rosenfield, S., Stansel, R. M., Bianchi, A., Moss, H., and de Lange, T. (1999) Mammalian telomeres end in a large duplex loop, *Cell 97*, 503-514.
85. Greider, C. W. (1999) Telomeres do d-loop-t-loop, *Cell 97*, 419-422.
86. de Lange, T. (2005) Shelterin: the protein complex that shapes and safeguards human telomeres, *Genes Dev 19*, 2100-2110.
87. Palm, W., and de Lange, T. (2008) How Shelterin Protects Mammalian Telomeres, *Annu Rev Genet 42*, 301-334.
88. Smogorzewska, A., and de Lange, T. (2004) Regulation of telomerase by telomeric proteins, *Annu Rev Biochem 73*, 177-208.
89. Shay, J. W. (1999) At the end of the millennium, a view of the end, *Nat Genet 23*, 382-383.
90. Ahmed, A., and Tollefsbol, T. (2003) Telomeres, Telomerase, and Telomerase Inhibition: Clinical Implications for Cancer, *Journal of the American Geriatrics Society 51*, 116-122.
91. Zhou, B. B., and Elledge, S. J. (2000) The DNA damage response: putting checkpoints in perspective, *Nature 408*, 433-439.
92. Zhu, X. D., Kuster, B., Mann, M., Petrini, J. H., and de Lange, T. (2000) Cell-cycle-regulated association of RAD50/MRE11/NBS1 with TRF2 and human telomeres, *Nat Genet 25*, 347-352.
93. Holt, S. E., and Shay, J. W. (1999) Role of telomerase in cellular proliferation and cancer, *J Cell Physiol 180*, 10-18.
94. Blackburn, E. H. (1992) Telomerases, *Annu Rev Biochem 61*, 113-129.
95. Wright, W. E., and Shay, J. W. (2001) Cellular senescence as a tumor-protection mechanism: the essential role of counting, *Curr Opin Genet Dev 11*, 98-103.
96. Holt, S. E., Aisner, D. L., Baur, J., Tesmer, V. M., Dy, M., Ouellette, M., Trager, J. B., Morin, G. B., Toft, D. O., Shay, J. W., Wright, W. E., and White, M. A. (1999) Functional requirement of p23 and Hsp90 in telomerase complexes, *Genes Dev 13*, 817-826.
97. Feng, J., Funk, W. D., Wang, S. S., Weinrich, S. L., Avilion, A. A., Chiu, C. P., Adams, R. R., Chang, E., Allsopp, R. C., Yu, J., and et al. (1995) The RNA component of human telomerase, *Science (New York, N.Y.) 269*, 1236-1241.
98. Cukusic, A., Skrobot Vidacek, N., Sopta, M., and Rubelj, I. (2008) Telomerase regulation at the crossroads of cell fate, *Cytogenetic and genome research 122*, 263-272.

99. Kyo, S., Takakura, M., Fujiwara, T., and Inoue, M. (2008) Understanding and exploiting hTERT promoter regulation for diagnosis and treatment of human cancers, *Cancer science 99*, 1528-1538.

100. Shaw, N. N., and Arya, D. P. (2008) Recognition of the unique structure of DNA:RNA hybrids, *Biochimie 90*, 1026-1039.

101. Rich, A. (1960) A hybrid helix containing both deoxyribose and ribose polynucleotides andits relation to the transfer of information between the nucleic acids, *Proceedings of the National Academy of Sciences of the United States of America 46*, 1044-1053.

102. Schildkraut, C. L., Marmur, J., and Doty, P. (1961) The formation of hybrid DNA molecules and their use in studies of DNA homologies, *Journal of molecular biology 3*, 595-IN516.

103. Gyi, J. I., Conn, G. L., Lane, A. N., and Brown, T. (1996) Comparison of the Thermodynamic Stabilities and Solution Conformations of DNA·RNA Hybrids Containing Purine-Rich and Pyrimidine-Rich Strands with DNA and RNA Duplexes, *Biochemistry 35*, 12538-12548.

104. Milman, G., Langridge, R., and Chamberlin, M. J. (1967) The structure of a DNA-RNA hybrid, *Proceedings of the National Academy of Sciences of the United States of America 57*, 1804-1810.

105. Gavory, G., Farrow, M., and Balasubramanian, S. (2002) Minimum length requirement of the alignment domain of human telomerase RNA to sustain catalytic activity in vitro, *Nucleic Acids Res 30*, 4470-4480.

106. Qi, X., Xie, M., Brown, A. F., Bley, C. J., Podlevsky, J. D., and Chen, J. J. (2012) RNA/DNA hybrid binding affinity determines telomerase template-translocation efficiency, *Embo J 31*, 150-161.

107. Francis, R., West, C., and Friedman, S. H. (2001) Targeting telomerase via its key RNA/DNA heteroduplex, *Bioorg Chem 29*, 107-117.

108. Martin, F. H., and Tinoco, I., Jr. (1980) DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription, *Nucleic Acids Res 8*, 2295-2299.

109. Temin, H. M. (1985) Reverse transcription in the eukaryotic genome: retroviruses, pararetroviruses, retrotransposons, and retrotranscripts, *Molecular Biology and Evolution 2*, 455-468.

110. Li, T. K., Barbieri, C. M., Lin, H. C., Rabson, A. B., Yang, G., Fan, Y., Gaffney, B. L., Jones, R. A., and Pilch, D. S. (2004) Drug targeting of HIV-1 RNA.DNA hybrid structures: thermodynamics of recognition and impact on reverse transcriptase-mediated ribonuclease H activity and viral replication, *Biochemistry 43*, 9732-9742.

111. Xu, B., and Clayton, D. A. (1996) RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA-DNA hybrids serving as primers, *The EMBO Journal 15*, 3135-3143.

112. Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K., and Sugino, A. (1968) Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains, *Proceedings of the National Academy of Sciences of the United States of America 59*, 598-605.

113. Chang, D. D., and Clayton, D. A. (1985) Priming of human mitochondrial DNA replication occurs at the light-strand promoter, *Proceedings of the National Academy of Sciences of the United States of America 82*, 351-355.

114. Lesnik, E. A., and Freier, S. M. (1995) Relative thermodynamic stability of DNA, RNA, and DNA:RNA hybrid duplexes: relationship with base composition and structure, *Biochemistry 34*, 10807-10815.

115. Roberts, R. W., and Crothers, D. M. (1992) Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition, *Science (New York, N.Y.) 258*, 1463-1466.

116. Egli, M., Usman, N., Zhang, S. G., and Rich, A. (1992) Crystal structure of an Okazaki fragment at 2-A resolution, *Proceedings of the National Academy of Sciences of the United States of America 89*, 534-538.

117. Egli, M., Usman, N., and Rich, A. (1993) Conformational influence of the ribose 2'-hydroxyl group: Crystal structures of DNA-RNA chimeric duplexes, *Biochemistry 32*, 3221-3237.

118. Gonzalez, C., Stec, W., Reynolds, M. A., and James, T. L. (1995) Structure and Dynamics of a DNA.cntdot.RNA Hybrid Duplex with a Chiral Phosphorothioate Moiety: NMR and Molecular Dynamics with Conventional and Time-Averaged Restraints, *Biochemistry 34*, 4969-4982.

119. Fedoroff, O. Y., Ge, Y., and Reid, B. R. (1997) Solution structure sf r(gaggacug):d(CAGTCCTC) hybrid: implications for the initiation of HIV-1(+)-strand synthesis1, *Journal of molecular biology 269*, 225-239.

120. Gyi, J. I., Lane, A. N., Conn, G. L., and Brown, T. (1998) Solution Structures of DNA·RNA Hybrids with Purine-Rich and Pyrimidine-Rich Strands: Comparison with the Homologous DNA and RNA Duplexes, *Biochemistry 37*, 73-80.

121. Ratmeyer, L., Vinayak, R., Zhong, Y. Y., Zon, G., and Wilson, W. D. (1994) Sequence Specific Thermodynamic and Structural Properties for DNA.RNA Duplexes, *Biochemistry 33*, 5298-5304.

122. Hung, S.-H., Yu, Q., Gray, D. M., and Ratliff, R. L. (1994) Evidence from CD spectra that d(purine)-r(pyrimidine) and r(purine)-d(pyrimidine) hybrids are in different structural classes, *Nucleic Acids Research 22*, 4326-4334.

123. Lesnik, E. A., and Freier, S. M. (1995) Relative Thermodynamic Stability of DNA, RNA, and DNA:RNA Hybrid Duplexes: Relationship with Base Composition and Structure, *Biochemistry 34*, 10807-10815.

124. Wang, S., and Kool, E. T. (1995) Origins of the Large Differences in Stability of DNA and RNA Helixes: C-5 Methyl and 2'-Hydroxyl Effects, *Biochemistry 34*, 4125-4132.

125. Milman, G., Langridge, R., and Chamberlin, M. J. (1967) The structure of a DNA-RNA hybrid, *Proceedings of the National Academy of Sciences of the United States of America 57*, 1804-1810.
126. Shindo, H., and Matsumoto, U. (1984) Direct evidence for a bimorphic structure of a DNA-RNA hybrid, poly(rA).poly(dT), at high relative humidity, *J Biol Chem 259*, 8682-8684.
127. Zimmerman, S. B., and Pheiffer, B. H. (1981) A RNA.DNA hybrid that can adopt two conformations: an x-ray diffraction study of poly(rA).poly(dT) in concentrated solution or in fibers, *Proceedings of the National Academy of Sciences of the United States of America 78*, 78-82.
128. Conn, G. L., Brown, T., and Leonard, G. A. (1999) The crystal structure of the RNA/DNA hybrid r(GAAGAGAAGC)·d(GCTTCTCTTC) shows significant differences to that found in solution, *Nucleic Acids Research 27*, 555-561.
129. Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I., Jr., and Blackburn, E. H. (1987) Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs, *Cell 51*, 899-908.
130. Sundquist, W. I., and Klug, A. (1989) Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops, *Nature 342*, 825-829.
131. Schaffitzel, C., Berger, I., Postberg, J., Hanes, J., Lipps, H. J., and Pluckthun, A. (2001) In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei, *Proceedings of the National Academy of Sciences of the United States of America 98*, 8572-8577.
132. Paeschke, K., Simonsson, T., Postberg, J., Rhodes, D., and Lipps, H. J. (2005) Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo, *Nature structural & molecular biology 12*, 847-854.
133. Yang, D. Z., and Okamoto, K. (2010) Structural insights into G-quadruplexes: towards new anticancer drugs, *Future Med Chem 2*, 619-646.
134. Lipps, H. J., and Rhodes, D. (2009) G-quadruplex structures: in vivo evidence and function, *Trends Cell Biol 19*, 414-422.
135. Zahler, A. M., Williamson, J. R., Cech, T. R., and Prescott, D. M. (1991) Inhibition of telomerase by G-quartet DNA structures, *Nature 350*, 718-720.
136. Oganesian, L., and Bryan, T. M. (2007) Physiological relevance of telomeric G-quadruplex formation: a potential drug target, *Bioessays 29*, 155-165.
137. Moon, I. K., and Jarstfer, M. B. (2007) The human telomere and its relationship to human disease, therapy, and tissue engineering, *Front Biosci 12*, 4595-4620.
138. Reed, J. E., Neidle, S., and Vilar, R. (2007) Stabilisation of human telomeric quadruplex DNA and inhibition of telomerase by a platinum-phenanthroline complex, *Chem Commun*, 4366-4368.
139. Neidle, S., and Parkinson, G. (2002) Telomere maintenance as a target for anticancer drug discovery, *Nature Reviews Drug Discovery 1*, 383-393.

140. Xu, Y., Suzuki, Y., Ishizuka, T., Xiao, C. D., Liu, X., Hayashi, T., and Komiyama, M. (2014) Finding a human telomere DNA-RNA hybrid G-quadruplex formed by human telomeric 6-mer RNA and 16-mer DNA using click chemistry: a protective structure for telomere end, *Bioorg Med Chem 22*, 4419-4421.
141. Phan, A. T., and Patel, D. J. (2003) Two-repeat human telomeric d(TAGGGTTAGGGT) sequence forms interconverting parallel and antiparallel G-quadruplexes in solution: distinct topologies, thermodynamic properties, and folding/unfolding kinetics, *J Am Chem Soc 125*, 15021-15027.
142. Phan, A. T., Luu, K. N., and Patel, D. J. (2006) Different loop arrangements of intramolecular human telomeric (3+1) G-quadruplexes in K(+) solution, *Nucleic Acids Research 34*, 5715-5719.
143. Marin, V. L., and Armitage, B. A. (2005) RNA guanine quadruplex invasion by complementary and homologous PNA probes, *J Am Chem Soc 127*, 8032-8033.
144. Wang, Y., and Patel, D. J. (1993) Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex, *Structure (London, England : 1993) 1*, 263-282.
145. Phan, A. T., Kuryavyi, V., Luu, K. N., and Patel, D. J. (2007) Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K+ solution, *Nucleic Acids Res 35*, 6517-6525.
146. Haider, S., Parkinson, G. N., and Neidle, S. (2002) Crystal structure of the potassium form of an Oxytricha nova G-quadruplex, *Journal of molecular biology 320*, 189-200.
147. Ambrus, A., Chen, D., Dai, J., Bialis, T., Jones, R. A., and Yang, D. (2006) Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution, *Nucleic Acids Research 34*, 2723-2735.
148. Luu, K. N., Phan, A. T., Kuryavyi, V., Lacroix, L., and Patel, D. J. (2006) Structure of the human telomere in K+ solution: an intramolecular (3 + 1) G-quadruplex scaffold, *J Am Chem Soc 128*, 9963-9970.
149. Todd, A. K., Johnston, M., and Neidle, S. (2005) Highly prevalent putative quadruplex sequence motifs in human DNA, *Nucleic Acids Res 33*, 2901-2907.
150. Huppert, J. L., and Balasubramanian, S. (2005) Prevalence of quadruplexes in the human genome, *Nucleic Acids Res 33*, 2908-2916.
151. Huppert, J. L., Bugaut, A., Kumari, S., and Balasubramanian, S. (2008) G-quadruplexes: the beginning and end of UTRs, *Nucleic Acids Res 36*, 6260-6268.
152. Neidle, S. (2009) The structures of quadruplex nucleic acids and their drug complexes, *Curr Opin Struct Biol 19*, 239-250.
153. Huppert, J. L. (2008) Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes, *Chem Soc Rev 37*, 1375-1384.
154. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K., and Neidle, S. (2006) Quadruplex DNA: sequence, topology and structure, *Nucleic Acids Research 34*, 5402-5415.

155. Wu, K. J., Grandori, C., Amacker, M., Simon-Vermot, N., Polack, A., Lingner, J., and Dalla-Favera, R. (1999) Direct activation of TERT transcription by c-MYC, *Nat Genet 21*, 220-224.

156. Yin, L., Hubbard, A. K., and Giardina, C. (2000) NF-kappa B regulates transcription of the mouse telomerase catalytic subunit, *J Biol Chem 275*, 36671-36675.

157. Lin, S. Y., and Elledge, S. J. (2003) Multiple tumor suppressor pathways negatively regulate telomerase, *Cell 113*, 881-889.

158. Hoffmeyer, K., Raggioli, A., Rudloff, S., Anton, R., Hierholzer, A., Del Valle, I., Hein, K., Vogt, R., and Kemler, R. (2012) Wnt/beta-catenin signaling regulates telomerase in stem cells and cancer cells, *Science (New York, N.Y.) 336*, 1549-1554.

159. Zhang, Y., Toh, L., Lau, P., and Wang, X. (2012) Human telomerase reverse transcriptase (hTERT) is a novel target of the Wnt/beta-catenin pathway in human cancer, *J Biol Chem 287*, 32494-32511.

160. Huang, F. W., Hodis, E., Xu, M. J., Kryukov, G. V., Chin, L., and Garraway, L. A. (2013) Highly recurrent TERT promoter mutations in human melanoma, *Science (New York, N.Y.) 339*, 957-959.

161. Horn, S., Figl, A., Rachakonda, P. S., Fischer, C., Sucker, A., Gast, A., Kadel, S., Moll, I., Nagore, E., Hemminki, K., Schadendorf, D., and Kumar, R. (2013) TERT promoter mutations in familial and sporadic melanoma, *Science (New York, N.Y.) 339*, 959-961.

162. Vinagre, J., Almeida, A., Populo, H., Batista, R., Lyra, J., Pinto, V., Coelho, R., Celestino, R., Prazeres, H., Lima, L., Melo, M., da Rocha, A. G., Preto, A., Castro, P., Castro, L., Pardal, F., Lopes, J. M., Santos, L. L., Reis, R. M., Cameselle-Teijeiro, J., Sobrinho-Simoes, M., Lima, J., Maximo, V., and Soares, P. (2013) Frequency of TERT promoter mutations in human cancers, *Nature communications 4*, 2185.

163. Killela, P. J., Reitman, Z. J., Jiao, Y., Bettegowda, C., Agrawal, N., Diaz, L. A., Jr., Friedman, A. H., Friedman, H., Gallia, G. L., Giovanella, B. C., Grollman, A. P., He, T. C., He, Y., Hruban, R. H., Jallo, G. I., Mandahl, N., Meeker, A. K., Mertens, F., Netto, G. J., Rasheed, B. A., Riggins, G. J., Rosenquist, T. A., Schiffman, M., Shih Ie, M., Theodorescu, D., Torbenson, M. S., Velculescu, V. E., Wang, T. L., Wentzensen, N., Wood, L. D., Zhang, M., McLendon, R. E., Bigner, D. D., Kinzler, K. W., Vogelstein, B., Papadopoulos, N., and Yan, H. (2013) TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal, *Proceedings of the National Academy of Sciences of the United States of America 110*, 6021-6026.

164. Allory, Y., Beukers, W., Sagrera, A., Flandez, M., Marques, M., Marquez, M., van der Keur, K. A., Dyrskjot, L., Lurkin, I., Vermeij, M., Carrato, A., Lloreta, J., Lorente, J. A., Carrillo-de Santa Pau, E., Masius, R. G., Kogevinas, M., Steyerberg, E. W., van Tilborg, A. A., Abas, C., Orntoft, T. F., Zuiverloon, T. C., Malats, N., Zwarthoff, E. C., and Real, F. X. (2014) Telomerase reverse transcriptase promoter mutations in

bladder cancer: high frequency across stages, detection in urine, and lack of association with outcome, *European urology 65*, 360-366.

165. Nagel, I., Szczepanowski, M., Martin-Subero, J. I., Harder, L., Akasaka, T., Ammerpohl, O., Callet-Bauchu, E., Gascoyne, R. D., Gesk, S., Horsman, D., Klapper, W., Majid, A., Martinez-Climent, J. A., Stilgenbauer, S., Tonnies, H., Dyer, M. J., and Siebert, R. (2010) Deregulation of the telomerase reverse transcriptase (TERT) gene by chromosomal translocations in B-cell malignancies, *Blood 116*, 1317-1320.
166. Reddel, R. R. (2010) Senescence: an antiviral defense that is tumor suppressive?, *Carcinogenesis 31*, 19-26.
167. Chen, Y. J., Campbell, H. G., Wiles, A. K., Eccles, M. R., Reddel, R. R., Braithwaite, A. W., and Royds, J. A. (2008) PAX8 regulates telomerase reverse transcriptase and telomerase RNA component in glioma, *Cancer Res 68*, 5724-5732.
168. Cao, Y., Bryan, T. M., and Reddel, R. R. (2008) Increased copy number of the TERT and TERC telomerase subunit genes in cancer cells, *Cancer science 99*, 1092-1099.
169. Bryan, T. M., Englezou, A., Gupta, J., Bacchetti, S., and Reddel, R. R. (1995) Telomere elongation in immortal human cells without detectable telomerase activity, *Embo J 14*, 4240-4248.
170. Greider, C. W., and Blackburn, E. H. (1989) A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis, *Nature 337*, 331-337.
171. Cesare, A. J., and Reddel, R. R. (2010) Alternative lengthening of telomeres: models, mechanisms and implications, *Nat Rev Genet 11*, 319-330.
172. Perrem, K., Bryan, T. M., Englezou, A., Hackl, T., Moy, E. L., and Reddel, R. R. (1999) Repression of an alternative mechanism for lengthening of telomeres in somatic cell hybrids, *Oncogene 18*, 3383-3390.
173. Ishii, Y., Tsuyama, N., Maeda, S., Tahara, H., and Ide, T. (1999) Telomerase activity in hybrids between telomerase-negative and telomerase-positive immortal human cells is repressed in the different complementation groups but not in the same complementation group of immortality, *Mechanisms of ageing and development 110*, 175-193.
174. Chung, I., Osterwald, S., Deeg, K. I., and Rippe, K. (2012) PML body meets telomere: the beginning of an ALTernate ending?, *Nucleus 3*, 263-275.
175. Vousden, K. H., and Lane, D. P. (2007) p53 in health and disease, *Nat Rev Mol Cell Biol 8*, 275-283.
176. Fagagna, F. d. A. d., Reaper, P. M., Clay-Farrace, L., Fiegler, H., Carr, P., von Zglinicki, T., Saretzki, G., Carter, N. P., and Jackson, S. P. (2003) A DNA damage checkpoint response in telomere-initiated senescence, *Nature 426*, 194-198.
177. Aubert, G., and Lansdorp, P. M. (2008) Telomeres and aging, *Physiol Rev 88*, 557-579.

178. Hahn, W. C., Counter, C. M., Lundberg, A. S., Beijersbergen, R. L., Brooks, M. W., and Weinberg, R. A. (1999) Creation of human tumour cells with defined genetic elements, *Nature 400*, 464-468.

179. Raymond, E., Sun, D., Chen, S. F., Windle, B., and VonHoff, D. D. (1996) Agents that target telomerase and telomeres, *Curr Opin Biotech 7*, 583-591.

180. Vonderheide, R. H. (2002) Telomerase as a universal tumor-associated antigen for cancer immunotherapy, *Oncogene 21*, 674-679.

181. Liu, J. P., Chen, W., Schwarer, A. P., and Li, H. (2010) Telomerase in cancer immunotherapy, *Biochim Biophys Acta 1805*, 35-42.

182. West, C., Francis, R., and Friedman, S. H. (2001) Small molecule/nucleic acid affinity chromatography: Application for the identification of telomerase inhibitors which target its key RNA/DNA heteroduplex, *Bioorg Med Chem Lett 11*, 2727-2730.

183. Gmeiner, W. H., Cui, W., Konerding, D. E., Keifer, P. A., Sharma, S. K., Soto, A. M., Marky, L. A., and Lown, J. W. (1999) Shape-selective recognition of a model Okazaki fragment by geometrically-constrained bis-distamycins, *J Biomol Struct Dyn 17*, 507-518.

184. Wheelhouse, R., and Chaires, J. (2010) Drug Binding to DNA·RNA Hybrid Structures, In *Drug-DNA Interaction Protocols* (Fox, K. R., Ed.), pp 55-70, Humana Press.

185. Ren, J., and Chaires, J. B. (1999) Sequence and structural selectivity of nucleic acid binding ligands, *Biochemistry 38*, 16067-16075.

186. Ren, J., Qu, X., Dattagupta, N., and Chaires, J. B. (2001) Molecular recognition of a RNA:DNA hybrid structure, *J Am Chem Soc 123*, 6742-6743.

187. Gillis, A. J., Schuller, A. P., and Skordalakes, E. (2008) Structure of the Tribolium castaneum telomerase catalytic subunit TERT, *Nature 455*, 633-637.

188. Rangarajan, S., and Friedman, S. H. (2007) Design, synthesis, and evaluation of phenanthridine derivatives targeting the telomerase RNA/DNA heteroduplex, *Bioorg Med Chem Lett 17*, 2267-2273.

189. Yu, H. Q., Zhang, D. H., Gu, X. B., Miyoshi, D., and Sugimoto, N. (2008) Regulation of telomerase activity by the thermodynamic stability of a DNA x RNA hybrid, *Angew Chem Int Ed Engl 47*, 9034-9038.

190. Barbieri, C. M., Li, T. K., Guo, S., Wang, G., Shallop, A. J., Pan, W., Yang, G., Gaffney, B. L., Jones, R. A., and Pilch, D. S. (2003) Aminoglycoside complexation with a DNA.RNA hybrid duplex: the thermodynamics of recognition and inhibition of RNA processing enzymes, *J Am Chem Soc 125*, 6469-6477.

191. Shaw, N. N., Xi, H., and Arya, D. P. (2008) Molecular recognition of a DNA:RNA hybrid: sub-nanomolar binding by a neomycin-methidium conjugate, *Bioorg Med Chem Lett 18*, 4142-4145.

192. Wheelhouse, R. T., Garbett, N. C., Buurma, N. J., and Chaires, J. B. (2010) Probing the Molecular Recognition of a DNA center dot RNA Hybrid Duplex, *Angew Chem Int Edit 49*, 3207-3210.

193. Campbell, N., and Neidle, S. (2012) G-Quadruplexes and Metal Ions, In *Interplay between Metal Ions and Nucleic Acids* (Sigel, A., Sigel, H., and Sigel, R. K. O., Eds.), pp 119-134, Springer Netherlands.
194. Yaku, H., Fujimoto, T., Murashima, T., Miyoshi, D., and Sugimoto, N. (2012) Phthalocyanines: a new class of G-quadruplex-ligands with many potential applications, *Chem Commun 48*, 6203-6216.
195. Alcaro, S., Artese, A., Costa, G., Distinto, S., Ortuso, F., and Parrotta, L. (2011) Conformational studies and solvent-accessible surface area analysis of known selective DNA G-Quadruplex binders, *Biochimie 93*, 1267-1274.
196. Parkinson, G. N., Lee, M. P., and Neidle, S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA, *Nature 417*, 876-880.
197. Dai, J., Punchihewa, C., Ambrus, A., Chen, D., Jones, R. A., and Yang, D. (2007) Structure of the intramolecular human telomeric G-quadruplex in potassium solution: a novel adenine triple formation, *Nucleic Acids Res 35*, 2440-2450.
198. Dai, J., Carver, M., Punchihewa, C., Jones, R. A., and Yang, D. (2007) Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K+ solution: insights into structure polymorphism of the human telomeric sequence, *Nucleic Acids Res 35*, 4927-4940.
199. Islam, M. K., Jackson, P. J., Rahman, K. M., and Thurston, D. E. (2016) Recent advances in targeting the telomeric G-quadruplex DNA sequence with small molecules as a strategy for anticancer therapies, *Future Med Chem 8*, 1259-1290.
200. Fechter, E. J., Olenyuk, B., and Dervan, P. B. (2004) Design of a Sequence-Specific DNA Bisintercalator, *Angewandte Chemie International Edition 43*, 3591-3594.
201. Sun, D., Thompson, B., Cathers, B. E., Salazar, M., Kerwin, S. M., Trent, J. O., Jenkins, T. C., Neidle, S., and Hurley, L. H. (1997) Inhibition of human telomerase by a G-quadruplex-interactive compound, *J. Med. Chem. 40*, 2113-2116.
202. Perry, P. J., Read, M. A., Davies, R. T., Gowan, S. M., Reszka, A. P., Wood, A. A., Kelland, L. R., and Neidle, S. (1999) 2,7-Disubstituted amidofluorenone derivatives as inhibitors of human telomerase, *J. Med. Chem. 42*, 2679-2684.
203. Read, M., Harrison, R. J., Romagnoli, B., Tanious, F. A., Gowan, S. H., Reszka, A. P., Wilson, W. D., Kelland, L. R., and Neidle, S. (2001) Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors, *Proceedings of the National Academy of Sciences of the United States of America 98*, 4844-4849.
204. Moore, M. J., Schultes, C. M., Cuesta, J., Cuenca, F., Gunaratnam, M., Tanious, F. A., Wilson, W. D., and Neidle, S. (2006) Trisubstituted acridines as G-quadruplex telomere targeting agents. Effects of extensions of the 3,6- and 9-side chains on quadruplex binding, telomerase activity, and cell proliferation, *J. Med. Chem. 49*, 582-599.
205. Burger, A. M., Dai, F., Schultes, C. M., Reszka, A. P., Moore, M. J., Double, J. A., and Neidle, S. (2005) The G-quadruplex-interactive molecule BRACO-19 inhibits tumor

growth, consistent with telomere targeting and interference with telomerase function, *Cancer Res 65*, 1489-1496.

206. Gunaratnam, M., Greciano, O., Martins, C., Reszka, A. P., Schultes, C. M., Morjani, H., Riou, J. F., and Neidle, S. (2007) Mechanism of acridine-based telomerase inhibition and telomere shortening, *Biochem Pharmacol 74*, 679-689.
207. Incles, C. M., Schultes, C. M., Kempski, H., Koehler, H., Kelland, L. R., and Neidle, S. (2004) A G-quadruplex telomere targeting agent produces p16-associated senescence and chromosomal fusions in human prostate cancer cells, *Molecular cancer therapeutics 3*, 1201-1206.
208. Riou, J. F., Guittat, L., Mailliet, P., Laoui, A., Renou, E., Petitgenet, O., Megnin-Chanet, F., Helene, C., and Mergny, J. L. (2002) Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands, *Proceedings of the National Academy of Sciences of the United States of America 99*, 2672-2677.
209. Sparapani, S., Haider, S. M., Doria, F., Gunaratnam, M., and Neidle, S. (2010) Rational design of acridine-based ligands with selectivity for human telomeric quadruplexes, *J Am Chem Soc 132*, 12263-12272.
210. Shin-ya, K., Wierzba, K., Matsuo, K., Ohtani, T., Yamada, Y., Furihata, K., Hayakawa, Y., and Seto, H. (2001) Telomestatin, a novel telomerase inhibitor from Streptomyces anulatus, *J Am Chem Soc 123*, 1262-1263.
211. Doi, T., Yoshida, M., Shin-ya, K., and Takahashi, T. (2006) Total Synthesis of (R)-Telomestatin, *Organic Letters 8*, 4165-4167.
212. Shammas, M. A., Shmookler Reis, R. J., Li, C., Koley, H., Hurley, L. H., Anderson, K. C., and Munshi, N. C. (2004) Telomerase inhibition and cell growth arrest after telomestatin treatment in multiple myeloma, *Clinical cancer research : an official journal of the American Association for Cancer Research 10*, 770-776.
213. Tauchi, T., Shin-Ya, K., Sashida, G., Sumi, M., Nakajima, A., Shimamoto, T., Ohyashiki, J. H., and Ohyashiki, K. (2003) Activity of a novel G-quadruplex-interactive telomerase inhibitor, telomestatin (SOT-095), against human leukemia cells: involvement of ATM-dependent DNA damage response pathways, *Oncogene 22*, 5338-5347.
214. Gomez, D., O'Donohue, M. F., Wenner, T., Douarre, C., Macadre, J., Koebel, P., Giraud-Panis, M. J., Kaplan, H., Kolkes, A., Shin-ya, K., and Riou, J. F. (2006) The G-quadruplex ligand telomestatin inhibits POT1 binding to telomeric sequences in vitro and induces GFP-POT1 dissociation from telomeres in human cells, *Cancer Res 66*, 6908-6912.
215. Tahara, H., Shin-Ya, K., Seimiya, H., Yamada, H., Tsuruo, T., and Ide, T. (2006) G-Quadruplex stabilization by telomestatin induces TRF2 protein dissociation from telomeres and anaphase bridge formation accompanied by loss of the 3' telomeric overhang in cancer cells, *Oncogene 25*, 1955-1966.
216. Izbicka, E., Nishioka, D., Marcell, V., Raymond, E., Davidson, K. K., Lawrence, R. A., Wheelhouse, R. T., Hurley, L. H., Wu, R. S., and Von Hoff, D. D. (1999) Telomere-

interactive agents affect proliferation rates and induce chromosomal destabilization in sea urchin embryos, *Anti-cancer drug design 14*, 355-365.

217. Grand, C. L., Han, H., Munoz, R. M., Weitman, S., Von Hoff, D. D., Hurley, L. H., and Bearss, D. J. (2002) The cationic porphyrin TMPyP4 down-regulates c-MYC and human telomerase reverse transcriptase expression and inhibits tumor growth in vivo, *Molecular cancer therapeutics 1*, 565-573.
218. Yang, D., and Okamoto, K. (2010) Structural insights into G-quadruplexes: towards new anticancer drugs, *Future Med Chem 2*, 619-646.
219. Gowan, S. M., Heald, R., Stevens, M. F., and Kelland, L. R. (2001) Potent inhibition of telomerase by small-molecule pentacyclic acridines capable of interacting with G-quadruplexes, *Molecular pharmacology 60*, 981-988.
220. Cookson, J. C., Dai, F., Smith, V., Heald, R. A., Laughton, C. A., Stevens, M. F., and Burger, A. M. (2005) Pharmacodynamics of the G-quadruplex-stabilizing telomerase inhibitor 3,11-difluoro-6,8,13-trimethyl-8H-quino[4,3,2-kl]acridinium methosulfate (RHPS4) in vitro: activity in human tumor cells correlates with telomere length and can be enhanced, or antagonized, with cytotoxic agents, *Molecular pharmacology 68*, 1551-1558.
221. Nakatani, K., Hagihara, S., Sando, S., Sakamoto, S., Yamaguchi, K., Maesawa, C., and Saito, I. (2003) Induction of a remarkable conformational change in a human telomeric sequence by the binding of naphthyridine dimer: inhibition of the elongation of a telomeric repeat by telomerase, *J Am Chem Soc 125*, 662-666.
222. Islam, M. M., Fujii, S., Sato, S., Okauchi, T., and Takenaka, S. (2015) A Selective G-Quadruplex DNA-Stabilizing Ligand Based on a Cyclic Naphthalene Diimide Derivative, *Molecules 20*, 10963-10979.
223. Lu, Y. J., Ou, T. M., Tan, J. H., Hou, J. Q., Shao, W. Y., Peng, D., Sun, N., Wang, X. D., Wu, W. B., Bu, X. Z., Huang, Z. S., Ma, D. L., Wong, K. Y., and Gu, L. Q. (2008) 5-N-methylated quindoline derivatives as telomeric g-quadruplex stabilizing ligands: effects of 5-N positive charge on quadruplex binding affinity and cell proliferation, *J. Med. Chem. 51*, 6381-6392.
224. Larsen, A. F., Nielsen, M. C., and Ulven, T. (2012) Tetrasubstituted phenanthrolines as highly potent, water-soluble, and selective G-quadruplex ligands, *Chemistry 18*, 10892-10902.
225. Tera, M., Hirokawa, T., Okabe, S., Sugahara, K., Seimiya, H., and Shimamoto, K. (2015) Design and Synthesis of a Berberine Dimer: A Fluorescent Ligand with High Affinity towards G-Quadruplexes, *Chemistry 21*, 14519-14528.
226. Bhowmik, D., Fiorillo, G., Lombardi, P., and Suresh Kumar, G. (2015) Recognition of human telomeric G-quadruplex DNA by berberine analogs: effect of substitution at the 9 and 13 positions of the isoquinoline moiety, *Journal of molecular recognition : JMR*.

227. Tan, J. H., Ou, T. M., Hou, J. Q., Lu, Y. J., Huang, S. L., Luo, H. B., Wu, J. Y., Huang, Z. S., Wong, K. Y., and Gu, L. Q. (2009) Isaindigotone derivatives: a new class of highly selective ligands for telomeric G-quadruplex DNA, *J. Med. Chem. 52*, 2825-2835.
228. Salem, A. A., El Haty, I. A., Abdou, I. M., and Mu, Y. (2015) Interaction of human telomeric G-quadruplex DNA with thymoquinone: a possible mechanism for thymoquinone anticancer effect, *Biochim Biophys Acta 1850*, 329-342.
229. Rahman, K. M., Hossain, M. D., Sohrab, M. H., Drake, A. F., Bui, T. T., Husby, J., Gunaratnam, M., Neidle, S., Hasan, C. M., and Thurston, D. E. (2012) The prenylated dioxopiperazine alkaloid Cristatin A has selective telomeric DNA G-quadruplex stabilising properties, *Chemical Communications 48*, 8760-8762.
230. Dhamodharan, V., Harikrishna, S., Bhasikuttan, A. C., and Pradeepkumar, P. I. (2015) Topology specific stabilization of promoter over telomeric G-quadruplex DNAs by bisbenzimidazole carboxamide derivatives, *ACS Chem Biol 10*, 821-833.
231. Kopel, P., Wawrzak, D., Langer, V., Cihalova, K., Chudobova, D., Vesely, R., Adam, V., and Kizek, R. (2015) Biological Activity and Molecular Structures of Bis(benzimidazole) and Trithiocyanurate Complexes, *Molecules 20*, 10360-10376.
232. Kralova, V., Hanusova, V., Stankova, P., Knoppova, K., Canova, K., and Skalova, L. (2013) Antiproliferative effect of benzimidazole anthelmintics albendazole, ricobendazole, and flubendazole in intestinal cancer cell lines, *Anti-cancer drugs 24*, 911-919.
233. Haque, R. A., Iqbal, M. A., Asekunowo, P., Majid, A. M. S. A., Ahamed, M. B. K., Umar, M. I., Al-Rawi, S. S., and Al-Suede, F. S. R. (2013) Synthesis, structure, anticancer, and antioxidant activity of para-xylyl linked bis-benzimidazolium salts and respective dinuclear Ag(I) N-heterocyclic carbene complexes (Part-II), *Med Chem Res 22*, 4663-4676.
234. Jain, A. K., Reddy, V. V., Paul, A., K, M., and Bhattacharya, S. (2009) Synthesis and evaluation of a novel class of G-quadruplex-stabilizing small molecules based on the 1,3-phenylene-bis(piperazinyl benzimidazole) system, *Biochemistry 48*, 10693-10704.
235. Paul, A., Jain, A. K., Misra, S. K., Maji, B., Muniyappa, K., and Bhattacharya, S. (2012) Binding of gemini bisbenzimidazole drugs with human telomeric G-quadruplex dimers: effect of the spacer in the design of potent telomerase inhibitors, *PLoS One 7*, e39467.
236. Maji, B., Kumar, K., Kaulage, M., Muniyappa, K., and Bhattacharya, S. (2014) Design and synthesis of new benzimidazole-carbazole conjugates for the stabilization of human telomeric DNA, telomerase inhibition, and their selective action on cancer cells, *J. Med. Chem. 57*, 6973-6988.
237. Drewe, W. C., Nanjunda, R., Gunaratnam, M., Beltran, M., Parkinson, G. N., Reszka, A. P., Wilson, W. D., and Neidle, S. (2008) Rational design of substituted diarylureas: a scaffold for binding to G-quadruplex motifs, *J. Med. Chem. 51*, 7751-7767.

238. Roth, A., Harley, C. B., and Baerlocher, G. M. (2010) Imetelstat (GRN163L)--telomerase-based cancer therapy, *Recent results in cancer research. Fortschritte der Krebsforschung. Progres dans les recherches sur le cancer 184*, 221-234.

239. Fong, P. C., Boss, D. S., Yap, T. A., Tutt, A., Wu, P., Mergui-Roelvink, M., Mortimer, P., Swaisland, H., Lau, A., O'Connor, M. J., Ashworth, A., Carmichael, J., Kaye, S. B., Schellens, J. H., and de Bono, J. S. (2009) Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers, *The New England journal of medicine 361*, 123-134.

240. Ruden, M., and Puri, N. (2013) Novel anticancer therapeutics targeting telomerase, *Cancer Treat Rev 39*, 444-456.

241. Goetz, M. P., Toft, D. O., Ames, M. M., and Erlichman, C. (2003) The Hsp90 chaperone complex as a novel target for cancer therapy, *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO 14*, 1169-1176.

242. van Steensel, B., Smogorzewska, A., and de Lange, T. (1998) TRF2 Protects Human Telomeres from End-to-End Fusions, *Cell 92*, 401-413.

243. Coleman, C., Levine, D., Kishore, R., Qin, G., Thorne, T., Lambers, E., Sasi, S. P., Yaar, M., Gilchrest, B. A., and Goukassian, D. A. (2010) Inhibition of melanoma angiogenesis by telomere homolog oligonucleotides, *Journal of oncology 2010*, 928628.

244. Lee-Bellantoni, M. S., Yaar, M., Eller, M. S., Rünger, T. M., Gao, Y., and Gilchrest, B. A. (2009) Telomeric DNA Induces p53-Dependent Reactive Oxygen Species And Protects Against Oxidative Damage, *Journal of dermatological science 56*, 154-162.

245. Damm, K., Hemmann, U., Garin-Chesa, P., Hauel, N., Kauffmann, I., Priepke, H., Niestroj, C., Daiber, C., Enenkel, B., Guilliard, B., Lauritsch, I., Muller, E., Pascolo, E., Sauter, G., Pantic, M., Martens, U. M., Wenz, C., Lingner, J., Kraut, N., Rettig, W. J., and Schnapp, A. (2001) A highly selective telomerase inhibitor limiting human cancer cell proliferation, *Embo J 20*, 6958-6968.

246. Belair, C. D., Yeager, T. R., Lopez, P. M., and Reznikoff, C. A. (1997) Telomerase activity: A biomarker of cell proliferation, not malignant transformation, *Proceedings of the National Academy of Sciences of the United States of America 94*, 13677-13682.

247. Skvortsov, D. A., Zvereva, M. E., Shpanchenko, O. V., and Dontsova, O. A. (2011) Assays for Detection of Telomerase Activity, *Acta Naturae 3*, 48-68.

248. Szatmari, I., Tokes, S., Dunn, C. B., Bardos, T. J., and Aradi, J. (2000) Modified telomeric repeat amplification protocol: a quantitative radioactive assay for telomerase without using electrophoresis, *Analytical biochemistry 282*, 80-88.

249. Gomez, D., Mergny, J. L., and Riou, J. F. (2002) Detection of telomerase inhibitors based on g-quadruplex ligands by a modified telomeric repeat amplification protocol assay, *Cancer Res 62*, 3365-3368.

250. Savoysky, E., Akamatsu, K., Tsuchiya, M., and Yamazaki, T. (1996) Detection of telomerase activity by combination of TRAP method and scintillation proximity assay (SPA), *Nucleic Acids Res 24*, 1175-1176.

251. Hirose, M., Abe-Hashimoto, J., Ogura, K., Tahara, H., Ide, T., and Yoshimura, T. (1997) A rapid, useful and quantitative method to measure telomerase activity by hybridization protection assay connected with a telomeric repeat amplification protocol, *Journal of cancer research and clinical oncology 123*, 337-344.

252. Wu, Y. Y., Hruszkewycz, A. M., Delgado, R. M., Yang, A., Vortmeyer, A. O., Moon, Y. W., Weil, R. J., Zhuang, Z., and Remaley, A. T. (2000) Limitations on the quantitative determination of telomerase activity by the electrophoretic and ELISA based TRAP assays, *Clinica chimica acta; international journal of clinical chemistry 293*, 199-212.

253. Hou, M., Xu, D., Bjorkholm, M., and Gruber, A. (2001) Real-time quantitative telomeric repeat amplification protocol assay for the detection of telomerase activity, *Clinical chemistry 47*, 519-524.

254. Ohyashiki, K., Yahata, N., Ohyashiki, J. H., Iwama, H., Hayashi, S., Ando, K., Aizawa, T., Ito, T., Miki, M., and Ebihara, Y. (1998) A combination of semiquantitative telomerase assay and in-cell telomerase activity measurement using exfoliated urothelial cells for the detection of urothelial neoplasia, *Cancer 83*, 2554-2560.

255. Blackburn, E. H., Greider, C. W., Henderson, E., Lee, M. S., Shampay, J., and Shippen-Lentz, D. (1989) Recognition and elongation of telomeres by telomerase, *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada 31*, 553-560.

256. Maesawa, C., Inaba, T., Sato, H., Iijima, S., Ishida, K., Terashima, M., Sato, R., Suzuki, M., Yashima, A., Ogasawara, S., Oikawa, H., Sato, N., Saito, K., and Masuda, T. (2003) A rapid biosensor chip assay for measuring of telomerase activity using surface plasmon resonance, *Nucleic Acids Res 31*, E4-4.

257. Schmidt, P. M., Matthes, E., Scheller, F. W., Bienert, M., Lehmann, C., Ehrlich, A., and Bier, F. F. (2002) Real-time determination of telomerase activity in cell extracts using an optical biosensor, *Biological chemistry 383*, 1659-1666.

258. Grimm, J., Perez, J. M., Josephson, L., and Weissleder, R. (2004) Novel nanosensors for rapid analysis of telomerase activity, *Cancer Res 64*, 639-643.

259. Pavlov, V., Willner, I., Dishon, A., and Kotler, M. (2004) Amplified detection of telomerase activity using electrochemical and quartz crystal microbalance measurements, *Biosensors & bioelectronics 20*, 1011-1021.

260. Sato, S., Kondo, H., Nojima, T., and Takenaka, S. (2005) Electrochemical telomerase assay with ferrocenylnaphthalene diimide as a tetraplex DNA-specific binder, *Analytical chemistry 77*, 7304-7309.

261. Rahman, K. M., Tizkova, K., Reszka, A. P., Neidle, S., and Thurston, D. E. (2012) Identification of novel telomeric G-quadruplex-targeting chemical scaffolds through screening of three NCI libraries, *Bioorg Med Chem Lett 22*, 3006-3010.

262. Sinha, B. K., Philen, R. M., Sato, R., and Cysyk, R. L. (1977) Synthsis and antitumor properties of bis(quinaldine) derivatives, *J. Med. Chem. 20*, 1528-1531.

263. St George, S., Bishop, J. V., Titus, R. G., and Selitrennikoff, C. P. (2006) Novel compounds active against Leishmania major, *Antimicrob Agents Chemother 50*, 474-479.

264. Todd, A. K., Haider, S. M., Parkinson, G. N., and Neidle, S. (2007) Sequence occurrence and structural uniqueness of a G-quadruplex in the human c-kit promoter, *Nucleic Acids Res 35*, 5799-5808.

265. Monchaud, D., and Teulade-Fichou, M. P. (2008) A hitchhiker's guide to G-quadruplex ligands, *Organic & biomolecular chemistry 6*, 627-636.

266. Ou, T. M., Lu, Y. J., Tan, J. H., Huang, Z. S., Wong, K. Y., and Gu, L. Q. (2008) G-quadruplexes: targets in anticancer drug design, *ChemMedChem 3*, 690-713.

267. Lim, K. W., Amrane, S., Bouaziz, S., Xu, W., Mu, Y., Patel, D. J., Luu, K. N., and Phan, A. T. (2009) Structure of the human telomere in K+ solution: a stable basket-type G-quadruplex with only two G-tetrad layers, *J Am Chem Soc 131*, 4301-4309.

268. Phan, A. T. (2010) Human telomeric G-quadruplex: structures of DNA and RNA sequences, *FEBS J 277*, 1107-1117.

269. Hammond, P. W., and Cech, T. R. (1998) Euplotes telomerase: evidence for limited base-pairing during primer elongation and dGTP as an effector of translocation, *Biochemistry 37*, 5162-5172.

270. Wei, D., Parkinson, G. N., Reszka, A. P., and Neidle, S. (2012) Crystal structure of a c-kit promoter quadruplex reveals the structural role of metal ions and water molecules in maintaining loop conformation, *Nucleic Acids Res 40*, 4691-4700.

271. Harrison, R. J., Gowan, S. M., Kelland, L. R., and Neidle, S. (1999) Human telomerase inhibition by substituted acridine derivatives, *Bioorg Med Chem Lett 9*, 2463-2468.

272. Song, A. M., Marik, J., and Lam, K. S. (2004) Solid-phase synthesis of 3-alkyl-2-arylamino-3,4-dihydroquinazolines, *Tetrahedron Lett 45*, 2727-2730.

273. Kopple, K. D., and Nitecki, D. E. (1962) Reactivity of Cyclic Peptides. II.cyclo-L-Tyrosyl-L-histidyl andcyclo-L-Tyrosyltriglycyl-L-histidylglycyl, *Journal of the American Chemical Society 84*, 4457-4464.

274. Sheehan, J. C., and Hess, G. P. (1955) A New Method of Forming Peptide Bonds, *Journal of the American Chemical Society 77*, 1067-1068.

275. Corbett, A. D., and Gleason, J. L. (2002) Preparation of active esters on solid support for aqueous-phase peptide couplings, *Tetrahedron Lett 43*, 1369-1372.

276. König, W., and Geiger, R. (1973) N-Hydroxyverbindungen als Katalysatoren für die Aminolyse aktivierter Ester, *Chemische Berichte 106*, 3626-3635.

277. Förster, T. (1948) Zwischenmolekulare Energiewanderung und Fluoreszenz (Intermolecular energy migration and fluorescence), *Annalen der Physik 437*, 55-75.

278. Didenko, V. V. (2001) DNA probes using fluorescence resonance energy transfer (FRET): Designs and applications, *Biotechniques 31*, 1106-+.

279. Remedios, C. G. d. (2001) Fluorescence Resonance Energy Transfer, In *eLS*, John Wiley & Sons, Ltd.
280. Clegg, R. (2006) The History of Fret, In *Reviews in Fluorescence 2006* (Geddes, C., and Lakowicz, J., Eds.), pp 1-45, Springer US.
281. De Cian, A., Guittat, L., Kaiser, M., Sacca, B., Amrane, S., Bourdoncle, A., Alberti, P., Teulade-Fichou, M. P., Lacroix, L., and Mergny, J. L. (2007) Fluorescence-based melting assays for studying quadruplex ligands, *Methods 42*, 183-195.
282. Simonsson, T., and Sjoback, R. (1999) DNA tetraplex formation studied with fluorescence resonance energy transfer, *J Biol Chem 274*, 17379-17383.
283. Mergny, J. L. (1999) Fluorescence energy transfer as a probe for tetraplex formation: the i-motif, *Biochemistry 38*, 1573-1581.
284. Mergny, J. L., and Maurizot, J. C. (2001) Fluorescence resonance energy transfer as a probe for G-quartet formation by a telomeric repeat, *Chembiochem 2*, 124-132.
285. Rachwal, P. A., and Fox, K. R. (2007) Quadruplex melting, *Methods 43*, 291-301.
286. Mergny, J. L., Lacroix, L., Teulade-Fichou, M. P., Hounsou, C., Guittat, L., Hoarau, M., Arimondo, P. B., Vigneron, J. P., Lehn, J. M., Riou, J. F., Garestier, T., and Helene, C. (2001) Telomerase inhibitors based on quadruplex ligands selected by a fluorescence assay, *Proceedings of the National Academy of Sciences of the United States of America 98*, 3062-3067.
287. Clegg, R. M. (1995) Fluorescence resonance energy transfer, *Curr Opin Biotechnol 6*, 103-110.
288. Mergny, J. L., Garestier, T., Rougee, M., Lebedev, A. V., Chassignol, M., Thuong, N. T., and Helene, C. (1994) Fluorescence energy transfer between two triple helix-forming oligonucleotides bound to duplex DNA, *Biochemistry 33*, 15321-15328.
289. Moore, G. E. (1998) Cramming more components onto integrated circuits (Reprinted from Electronics, pg 114-117, April 19, 1965), *P Ieee 86*, 82-85.
290. Shaw, D. E., Maragakis, P., Lindorff-Larsen, K., Piana, S., Dror, R. O., Eastwood, M. P., Bank, J. A., Jumper, J. M., Salmon, J. K., Shan, Y., and Wriggers, W. (2010) Atomic-level characterization of the structural dynamics of proteins, *Science (New York, N.Y.) 330*, 341-346.
291. Shaw, D. E., Deneroff, M. M., Dror, R. O., Kuskin, J. S., Larson, R. H., Salmon, J. K., Young, C., Batson, B., Bowers, K. J., Chao, J. C., Eastwood, M. P., Gagliardo, J., Grossman, J. P., Ho, C. R., Ierardi, D. J., Istv, #225, Kolossv, n., ry, Klepeis, J. L., Layman, T., McLeavey, C., Moraes, M. A., Mueller, R., Priest, E. C., Shan, Y., Spengler, J., Theobald, M., Towles, B., and Wang, S. C. (2007) Anton, a special-purpose machine for molecular dynamics simulation, In *Proceedings of the 34th annual international symposium on Computer architecture*, pp 1-12, ACM, San Diego, California, USA.
292. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics, *J Mol Graph 14*, 33-38, 27-38.

293. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis, *J Comput Chem 25*, 1605-1612.

294. Ranjbar, B., and Gill, P. (2009) Circular dichroism techniques: biomolecular and nanostructural analyses- a review, *Chemical biology & drug design 74*, 101-120.

295. Sun, X. G., Cao, E. H., He, Y. J., and Qin, J. F. (1999) Spectroscopic comparison of different DNA structures formed by oligonucleotides, *J Biomol Struct Dyn 16*, 863-872.

296. Kypr, J., Kejnovská, I., Renčiuk, D., and Vorlíčková, M. (2009) Circular dichroism and conformational polymorphism of DNA, *Nucleic Acids Research 37*, 1713-1725.

297. Xu, Y., Ishizuka, T., Yang, J., Ito, K., Katada, H., Komiyama, M., and Hayashi, T. (2012) Oligonucleotide models of telomeric DNA and RNA form a Hybrid G-quadruplex structure as a potential component of telomeres, *J Biol Chem 287*, 41787-41796.

298. Marck, C., and Thiele, D. (1978) Poly(dG).poly(dC) at neutral and alkaline pH: the formation of triple stranded poly(dG).poly(dG).poly(dC), *Nucleic Acids Research 5*, 1017-1028.

299. Zhang Xiaoyan, Cao Enhua, Sun Xueguang, and Chunl, B. (2000) Circular dichroism spectroscopic studies on structures formed by telomeric DNA sequences in vitro, *Chinese Science Bulletin 45*.

300. Balagurumoorthy, P., Brahmachari, S. K., Mohanty, D., Bansal, M., and Sasisekharan, V. (1992) Hairpin and parallel quartet structures for telomeric sequences, *Nucleic Acids Res 20*, 4061-4067.

301. Wang, Y., and Patel, D. J. (1992) Guanine residues in d(T2AG3) and d(T2G4) form parallel-stranded potassium cation stabilized G-quadruplexes with anti glycosidic torsion angles in solution, *Biochemistry 31*, 8112-8119.

302. Williamson, J. R. (1994) G-quartet structures in telomeric DNA, *Annu Rev Biophys Biomol Struct 23*, 703-730.

303. Berridge, M. V., Herst, P. M., and Tan, A. S. (2005) Tetrazolium dyes as tools in cell biology: new insights into their cellular reduction, *Biotechnology annual review 11*, 127-152.

304. Mosmann, T. (1983) Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays, *Journal of immunological methods 65*, 55-63.

305. Bernas, T., and Dobrucki, J. (2002) Mitochondrial and nonmitochondrial reduction of MTT: interaction of MTT with TMRE, JC-1, and NAO mitochondrial fluorescent probes, *Cytometry 47*, 236-242.

306. Berridge, M. V., and Tan, A. S. (1993) Characterization of the cellular reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT): subcellular localization, substrate dependence, and involvement of mitochondrial electron transport in MTT reduction, *Archives of biochemistry and biophysics 303*, 474-482.

307. Jennerwein, M. M., and Eastman, A. (1991) A polymerase chain reaction-based method to detect cisplatin adducts in specific genes, *Nucleic Acids Res 19*, 6209-6214.

308. Salonga, D., Danenberg, K., Grem, J., Park, J., and Danenberg, P. (2002) Relative Gene Expression in Normal and Tumor Tissue by Quantitative RT-PCR, In *Telomeres and Telomerase* (Double, J., and Thompson, M., Eds.), pp 83-98, Humana Press.

309. Emrich, T., Chang, S.-Y., Karl, G., Panzinger, B., and Santini, C. (2002) Quantitative Detection of Telomerase Components by Real-Time, Online RT-PCR Analysis with the LightCycler, In *Telomeres and Telomerase* (Double, J., and Thompson, M., Eds.), pp 99-108, Humana Press.

310. Zhou, X., and Xing, D. (2012) Assays for human telomerase activity: progress and prospects, *Chem Soc Rev 41*, 4643-4656.

311. Weise, J. M., and Gunes, C. (2006) Telomeres and telomerase. A survey about methods and recent advances in cancer diagnostic and therapy, *Histology and histopathology 21*, 1249-1261.

312. Sun, P. M., Wei, L. H., Luo, M. Y., Liu, G., Wang, J. L., Mustea, A., Konsgen, D., Lichtenegger, W., and Sehouli, J. (2007) The telomerase activity and expression of hTERT gene can serve as indicators in the anti-cancer treatment of human ovarian cancer, *Eur J Obstet Gynecol Reprod Biol 130*, 249-257.

313. Shimojima, M., Komine, F., Hisatomi, H., Shimizu, T., Moriyama, M., and Arakawa, Y. (2004) Detection of telomerase activity, telomerase RNA component, and telomerase reverse transcriptase in human hepatocellular carcinoma, *Hepatol Res 29*, 31-38.

314. Shay, J. W., and Wright, W. E. (2011) Role of telomeres and telomerase in cancer, *Semin Cancer Biol 21*, 349-353.

315. Tchirkov, A., Rolhion, C., Kemeny, J. L., Irthum, B., Puget, S., Khalil, T., Chinot, O., Kwiatkowski, F., Perissel, B., Vago, P., and Verrelle, P. (2003) Clinical implications of quantitative real-time RT-PCR analysis of hTERT gene expression in human gliomas, *British journal of cancer 88*, 516-520.

316. Wasner, M., Tschop, K., Spiesbach, K., Haugwitz, U., Johne, C., Mossner, J., Mantovani, R., and Engeland, K. (2003) Cyclin B1 transcription is enhanced by the p300 coactivator and regulated during the cell cycle by a CHR-dependent repression mechanism, *FEBS letters 536*, 66-70.

317. Ausubel, F., M. . (1994) Current protocols in molecular biology. Chichester: John Wiley & Sons, 12.2.1-12.2.11. ISBN 0-471-50337-1, .

318. Lipinski, C. A. (2004) Lead- and drug-like compounds: the rule-of-five revolution, *Drug Discov Today Technol 1*, 337-341.

319. Ranjith, P. K., Rajeesh, P., Haridas, K. R., Susanta, N. K., Row, T. N., Rishikesan, R., and Kumari, N. S. (2013) Design and synthesis of positional isomers of 5 and 6-bromo-1-[(phenyl)sulfonyl]-2-[(4-nitrophenoxy)methyl]-1H-benzimidazoles as possible antimicrobial and antitubercular agents, *Bioorg Med Chem Lett 23*, 5228-5234.

320. Toro, P., Klahn, A. H., Pradines, B., Lahoz, F., Pascual, A., Biot, C., and Arancibia, R. (2013) Organometallic benzimidazoles: Synthesis, characterization and antimalarial activity, *Inorg Chem Commun 35*, 126-129.

321. Sondhi, S. M., Rajvanshi, S., Johar, M., Bharti, N., Azam, A., and Singh, A. K. (2002) Anti-inflammatory, analgesic and antiamoebic activity evaluation of pyrimido[1,6-a]benzimidazole derivatives synthesized by the reaction of ketoisothiocyanates with mono and diamines, *Eur J Med Chem 37*, 835-843.

322. Jain, A., Sharma, R., and Chaturvedi, S. C. (2013) A rational design, synthesis, characterization, and antihypertensive activities of some new substituted benzimidazoles, *Med Chem Res 22*, 4622-4632.

323. Beaulieu, C., Wang, Z., Denis, D., Greig, G., Lamontagne, S., O'Neill, G., Slipetz, D., and Wang, J. (2004) Benzimidazoles as new potent and selective DP antagonists for the treatment of allergic rhinitis, *Bioorg Med Chem Lett 14*, 3195-3199.

324. Evans, T. M., Gardiner, J. M., Mahmood, N., and Smis, M. (1997) Structure-activity relationships of anti-HIV-1 N-alkoxy- and N-allyloxy-benzimidazoles, *Bioorg Med Chem Lett 7*, 409-412.

325. Nile, S. H., Kumar, B., and Park, S. W. (2013) In vitro evaluation of selected benzimidazole derivatives as an antioxidant and xanthine oxidase inhibitors, *Chemical biology & drug design 82*, 290-295.

326. White, A. W., Curtin, N. J., Eastman, B. W., Golding, B. T., Hostomsky, Z., Kyle, S., Li, J., Maegley, K. A., Skalitzky, D. J., Webber, S. E., Yu, X. H., and Griffin, R. J. (2004) Potentiation of cytotoxic drug activity in human tumour cell lines, by amine-substituted 2-arylbenzimidazole-4-carboxamide PARP-1 inhibitors, *Bioorg Med Chem Lett 14*, 2433-2437.

327. Kahveci, B., Mentes, E., Ozil, M., Ulker, S., and Erturk, M. (2013) An efficient synthesis of benzimidazoles via a microwave technique and evaluation of their biological activities, *Monatsh Chem 144*, 993-1001.

328. Purushottamachar, P., Godbole, A. M., Gediya, L. K., Martin, M. S., Vasaitis, T. S., Kwegyir-Afful, A. K., Ramalingam, S., Ates-Alagoz, Z., and Njar, V. C. (2013) Systematic structure modifications of multitarget prostate cancer drug candidate galeterone to produce novel androgen receptor down-regulating agents as an approach to treatment of advanced prostate cancer, *J. Med. Chem. 56*, 4880-4898.

329. Case, D. A., Darden, T. A., Cheatham III, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., Kollman, P. A. (2010) AMBER 11, University of California, San Francisco, 2010.

330. Nunn, C. M., Jenkins, T. C., and Neidle, S. (1993) Crystal structure of d(CGCGAATTCGCG) complexed with propamidine, a short-chain homologue of the drug pentamidine, *Biochemistry 32*, 13838-13843.

331. Edwards, K. J., Jenkins, T. C., and Neidle, S. (1992) Crystal structure of a pentamidine-oligonucleotide complex: implications for DNA-binding properties, *Biochemistry 31*, 7104-7109.

332. Berman, H. M., Neidle, S., Zimmer, C., and Thrum, H. (1979) Netropsin, a DNA-binding oligopeptide structural and binding studies, *Biochim Biophys Acta 561*, 124-131.

333. Parkinson, G. N., Cuenca, F., and Neidle, S. (2008) Topology conservation and loop flexibility in quadruplex-drug recognition: crystal structures of inter- and intramolecular telomeric DNA quadruplex-drug complexes, *Journal of molecular biology 381*, 1145-1156.

334. Hung, S. H., Yu, Q., Gray, D. M., and Ratliff, R. L. (1994) Evidence from Cd-Spectra That D(Purine)Center-Dot-R(Pyrimidine) and R(Purine)Center-Dot-D(Pyrimidine) Hybrids Are in Different Structural Classes, *Nucleic Acids Research 22*, 4326-4334.

335. Vorlickova, M., Kejnovska, I., Bednarova, K., Renciuk, D., and Kypr, J. (2012) Circular dichroism spectroscopy of DNA: from duplexes to quadruplexes, *Chirality 24*, 691-698.

336. Balagurumoorthy, P., and Brahmachari, S. K. (1994) Structure and stability of human telomeric sequence, *J Biol Chem 269*, 21858-21869.

337. Paramasivan, S., Rujan, I., and Bolton, P. H. (2007) Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding, *Methods 43*, 324-331.

338. Nicoludis, J. M., Barrett, S. P., Mergny, J.-L., and Yatsunyk, L. A. (2012) Interaction of human telomeric DNA with N-methyl mesoporphyrin IX, *Nucleic Acids Research*.

339. 2015]., C. F. G. I. h. w. g. c. g. p. c.-f. i. h. r. d. t. h. c. J. ( ).

340. Jagadeesh, S., Kyo, S., and Banerjee, P. P. (2006) Genistein represses telomerase activity via both transcriptional and posttranslational mechanisms in human prostate cancer cells, *Cancer Res 66*, 2107-2115.

341. Yeo, M., Rha, S. Y., Jeung, H. C., Hu, S. X., Yang, S. H., Kim, Y. S., An, S. W., and Chung, H. C. (2005) Attenuation of telomerase activity by hammerhead ribozyme targeting human telomerase RNA induces growth retardation and apoptosis in human breast tumor cells, *International journal of cancer. Journal international du cancer 114*, 484-489.

342. Nakamura, T. M., Morin, G. B., Chapman, K. B., Weinrich, S. L., Andrews, W. H., Lingner, J., Harley, C. B., and Cech, T. R. (1997) Telomerase catalytic subunit homologs from fission yeast and human, *Science (New York, N.Y.) 277*, 955-959.

343. Wang, J., Xie, L. Y., Allan, S., Beach, D., and Hannon, G. J. (1998) Myc activates telomerase, *Genes Dev 12*, 1769-1774.

344. Kang, S. S., Kwon, T., Kwon, D. Y., and Do, S. I. (1999) Akt protein kinase enhances human telomerase activity through phosphorylation of telomerase reverse transcriptase subunit, *J Biol Chem 274*, 13085-13090.

345. Kyo, S., Kanaya, T., Takakura, M., Tanaka, M., and Inoue, M. (1999) Human telomerase reverse transcriptase as a critical determinant of telomerase activity in normal and malignant endometrial tissues, *International Journal of Cancer 80*, 60-63.

346. Greider, C. W. (1998) Telomerase activity, cell proliferation, and cancer, *Proceedings of the National Academy of Sciences of the United States of America 95*, 90-92.

347. Kirkpatrick, K. L., Clark, G., Ghilchick, M., Newbold, R. F., and Mokbel, K. (2003) hTERT mRNA expression correlates with telomerase activity in human breast cancer, *Eur J Surg Oncol 29*, 321-326.

348. Ito, M. (2000) Factors controlling cyclin B expression, *Plant molecular biology 43*, 677-690.

349. Hershko, A. (1999) Mechanisms and regulation of the degradation of cyclin B, *Philosophical transactions of the Royal Society of London. Series B, Biological sciences 354*, 1571-1575; discussion 1575-1576.

350. Androic, I., Kramer, A., Yan, R., Rodel, F., Gatje, R., Kaufmann, M., Strebhardt, K., and Yuan, J. (2008) Targeting cyclin B1 inhibits proliferation and sensitizes breast cancer cells to taxol, *BMC Cancer 8*, 391.

351. Aaltonen, K., Amini, R. M., Heikkila, P., Aittomaki, K., Tamminen, A., Nevanlinna, H., and Blomqvist, C. (2009) High cyclin B1 expression is associated with poor survival in breast cancer, *British journal of cancer 100*, 1055-1060.

352. Begnami, M. D., Fregnani, J. H., Nonogaki, S., and Soares, F. A. (2010) Evaluation of cell cycle protein expression in gastric cancer: cyclin B1 expression and its prognostic implication, *Human pathology 41*, 1120-1127.

353. McKerlie, M., and Zhu, X. D. (2011) Cyclin B-dependent kinase 1 regulates human TRF1 to modulate the resolution of sister telomeres, *Nature communications 2*, 371.

354. Kim, S., Lee, H. S., Lee, S. K., Kim, S. H., Hur, S. M., Kim, J. S., Kim, J. H., Choe, J. H., Shin, I., Yang, J. H., Lee, J. E., and Nam, S. J. (2010) 12-O-Tetradecanoyl phorbol-13-acetate (TPA)-induced growth arrest is increased by silibinin by the down-regulation of cyclin B1 and cdc2 and the up-regulation of p21 expression in MDA-MB231 human breast cancer cells, *Phytomedicine : international journal of phytotherapy and phytopharmacology 17*, 1127-1132.

355. Fried, M. G. (1989) Measurement of protein-DNA interaction parameters by electrophoresis mobility shift assay, *Electrophoresis 10*, 366-376.